

**LINKING RESOURCE ABUNDANCE AND ENVIRONMENTAL STOCHASTICITY  
WITH ANIMAL SPACE USE USING CONTINUOUS-TIME STOCHASTIC PROCESSES**

by

Stefano Mezzini

B.Sc. Hons. Biology, University of Regina, 2021

B.Sc. Statistics, University of Regina, 2021

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

May 2022

© Stefano Mezzini, 2022

The following individuals certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis entitled:

*Linking resource abundance and environmental stochasticity with animal space use using continuous-time stochastic processes,*

submitted by Stefano Mezzini in partial fulfillment of the requirements of the degree of Master of Science.

**Michael J. Noonan** \_\_\_\_\_

Supervisor

**Adam T. Ford** \_\_\_\_\_

Supervisory Committee Member

**Jason Pither** \_\_\_\_\_

Supervisory Committee Member

## **Abstract**

The ability to move allows animals to access more resources and decrease the risk of mortality or injury from various threats, including predation and competition. However, most motile animals depend on movement to achieve an energetic balance, survive, and reproduce. Animals' spatial needs depend on a multitude of factors, including resource abundance, competitive pressure, predation, weather, and climate. The effects of many of these factors have been studied extensively, but the effects of environmental heterogeneity and unpredictability remain largely understudied. This thesis aims to quantify how resource abundance and environmental variation affect animals' spatial needs using theoretical arguments, computational simulations, and a large animal movement dataset of over 1500 individuals from more than 75 species. To this end, this thesis will produce a new global measure of environmental variance which will be used to quantify the effect of environmental heterogeneity and stochasticity on animals' spatial needs while accounting for spatiotemporal trends and correlations within populations and species. The animals' movement will be modeled using continuous-time stochastic movement models. The effects of environmental variance on animal's spatial needs will be estimated using hierarchical generalized additive models that account for common trends between taxonomic groups and populations. The work will be carried out within a framework that recognizes the value of different forms of knowledge, including Traditional Indigenous Knowledge, and it will incorporate Traditional Indigenous Knowledge when possible. Overall, I expect spatial needs to be lower in areas with higher resource abundance and lower environmental variance. However, I also expect the effects of resource abundance and environmental variance to depend on the species and ecosystems of interest. Thus, the average trends between species may not be representative of individual populations and species.

## Lay Summary

Many factors affect how much space animals need to survive and reproduce. The effects of many factors have been studied extensively, such as resource abundance, competitive pressure, predation, weather, and climate. However, the effects of environmental heterogeneity and unpredictability remain understudied. This thesis aims to quantify how resource abundance and environmental variation affect animals' spatial needs using a large animal movement dataset. This thesis will produce a global measure of environmental variance which will be used to quantify the effect of such variance on animals' spatial needs while accounting for spatiotemporal trends and correlations within populations and species. The work will be carried out while recognizing the value of different forms of knowledge, including Traditional Indigenous Knowledge. I expect spatial needs to be lower in areas with higher resource abundance and lower environmental variance, although this will likely depend on the species and ecosystems of interest.

## Foreword

### Two-eyed seeing: Recognizing Traditional Indigenous Knowledge

The lands managed and protected by Indigenous Peoples are often markedly different from those inhabited by urban societies. While recognizing that there is great diversity between Indigenous Peoples (as well as other colonized Peoples), it is important to recognize that many hold great knowledge on how to live sustainably, safeguard environments, and protect biodiversity (Schuster *et al.*, 2019), and have been doing so for millennia. Yet, their leaders and representatives are seldom included in conservation-related decision-making. Instead, many Western institutions often dismiss, ignore, and contradict the ancestral and traditional Knowledge of Indigenous and colonized Peoples (Kimmerer, 2020; Smith, 2021). The development of Western science is frequently assumed to clash with the (often sacred) Knowledge many colonized Peoples hold. Western science is often viewed as more objective, methodical, and unbiased than traditional Knowledge, and as such Western institutions and people often consider it to be superior to Indigenous Knowledge (Smith, 2021). However, it is common for Western institutions to (reluctantly) reach similar, if not identical, conclusions as those held by Indigenous people (Kimmerer, 2020; Smith, 2021; Bennett *et al.*, 2021). The refusal to recognize traditional Knowledge and cooperate with non-Western institutions often results in a loss of time, resources, and funds to the Western institutions and severe damage to the Land the institution operated on, as well as to the people who's ancestors inhabited the region for millennia (Smith, 2021). The development of Western science at the exclusion of Indigenous Peoples perpetuates colonialism and brings harm all parties involved, and reinforces “research” as the dirtiest word for many Indigenous Peoples (Smith, 2021, 1). The sharing, development, recognition, and ownership (or lack thereof) of knowledge (and Sacred Knowledge) are core aspects of imperialism (Smith, 2021, xii).

The concept of *two-eye seeing* refers to an approach to knowledge and growth that braids Indigenous Knowledge and science together with Western science (Kutz & Tomaselli, 2019; Kimmerer, 2020). Since Traditional Indigenous Knowledge tends to be qualitative, while biological sciences tend to be quantitative, connecting the two is not always simple (Bowles *et al.*, 2021). One possibility, however, is to use Traditional Knowledge to create properly informed Bayesian priors (Bolstad & Curran, 2017). The validity of the priors can be ensured using prior predictive modeling (McElreath, 2016) to select priors that align with the Traditional Knowledge. This practice is not new (Girondot & Rizzo, 2015; Bélisle *et al.*, 2018), but it is rarely used, despite it aligning well with the philosophy of Bayesian statistics.

Although this project does not focus on Indigenous Knowledge or Data, I intend to carryout the work with an anti-imperial and anti-colonial stance, to combat problematic and questionable practices and views (rather

than risking perpetuating them). I recognize my ignorance in the field in the hopes of receiving valuable critiques, direction, and suggestions. During the project, I intend connecting with local Indigenous groups and Nations (including Westbank First Nation) to offer help in related fields and exchange knowledge (while recognizing that I do not have the right to all knowledge and that many forms and sources of Knowledge are sacred). Regardless of whether collaborations will be possible, I will actively work on removing any form of colonial bias in my work and perspectives while also combating colonialism, imperialism, and other forms of oppression.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Lay Summary</b>	<b>iv</b>
<b>Foreword</b>	<b>v</b>
Two-eyed seeing: Recognizing Traditional Indigenous Knowledge . . . . .	v
<b>Acknowledgements</b>	<b>xiii</b>
<b>Dedication</b>	<b>xiv</b>
<b>1 Chapter 1: Literature review and quantitative predictions</b>	<b>1</b>
1.1 Home ranges as proxies for animals' needs . . . . .	1
1.2 Effects of resource availability and productivity on spatial needs . . . . .	5
1.3 Effects of environmental variance on spatial needs . . . . .	6
1.3.1 The temporal scale of stochastic events . . . . .	7
1.3.2 Spatial needs in stochastic environments . . . . .	10
1.4 Interaction effects of resource availability and stochasticity . . . . .	12
1.5 Thesis structure and aims . . . . .	13
<b>2 Chapter 2: A new measure of environmental variance</b>	<b>15</b>
2.1 Quantifying environmental variance . . . . .	17
<b>3 Chapter 3: Movement analyses</b>	<b>19</b>
<b>4 Chapter 4: Synthesis</b>	<b>22</b>
<b>5 Open and transparent science: Code and data availability</b>	<b>23</b>
<b>6 List of abbreviations used</b>	<b>24</b>
<b>7 List of notations and symbols used</b>	<b>25</b>
<b>8 Project timeline</b>	<b>26</b>
8.1 Progress to date . . . . .	26
<b>Bibliography</b>	<b>27</b>



## **List of Tables**

1	A list of all abbreviations and acronyms used in the thesis. . . . .	24
2	A list of all mathematical notations and symbols used in the thesis. . . . .	25

## List of Figures

1	Inset: Simulated movement of an animal with a stationary home range centroid. The colored areas indicate the home range estimates for different quantiles. The red area indicates the core home range (50%), while the blue and green areas indicate the 75% and 95% home ranges, respectively. Main figure: Home range area as a function of utilization quantile. Note that as the quantile approaches 1, home range size approaches infinity, since if the animal was to move for an infinite amount of time, it would cover all possible space. . . . .	2
2	Movement of a tapir during a monitoring period of approximately 3 years (1997-07-10 to 2000-06-08) used in the work by Medici <i>et al.</i> (2022). Different subsamples may result in different estimates of home range size and centroid (a, b; each span 100 days), so neither subsample is necessarily representative of the space the animal used over the entire monitoring period (c).	3
3	Fictitious example of variation in resource abundance in a heterogeneous but fully predictable environment. (a.) Complete predictability results in guaranteed successes (dots) during each foraging attempt. (b.) Arbitrary definition of $R$ as following a Gamma distribution with shape $k$ and scale $\theta$ , while $S$ follows a Bernoulli distribution with probability of success $p = 1$ , since successes always occur (i.e., $P(S = 0) = 0$ implies $S = 1$ ). Thus, $U = R \cdot S = R$ , so $U$ follows the same distribution as $R$ . The shaded green area indicates the probability of obtaining an arbitrarily sufficient amount of resources to survive and reproduce. (c.) Since foraging attempts are guaranteed to be successful, the amount of usable resources will equal the resource abundance. (d.) Since $U = R$ , the expected (i.e. average) amount of usable resources is equal to the expected resource abundance. . . . .	5
4	Fictitious changes in resource abundance mean (a) and variance (b) over the course of a year. The true trends are indicated by the black line, while the colored lines indicate the estimates perceived based on different levels of adaptability. The dashed vertical line indicates the hypothetical current date, such that any data to the right of the line is unknown to the animals, so estimates on the right of the dashed line are extrapolations. The data were simulated using Gaussian noise to ensure mean-variance independence and modeled using Generalized Additive Models with an identity link function to allow linear responses. . . . .	9

5	Fictitious example of variation in resource abundance in a heterogeneous and stochastic environment. (a.) Foraging successes (dots) can vary due to various reasons, including differences in competition, predation, and resource-specific trends. Successes most stochastic when the probability of success is approximately 0.5, while they are least stochastic (i.e., most predictable) when they are guaranteed to occur ( $p = 1$ ) or guaranteed to fail ( $p = 0$ ). (b.) Arbitrary definition of $R$ as following a Gamma distribution with shape $k$ and scale $\theta$ , while $S$ follows a Bernoulli distribution with probability of success $p = 0.6$ , so successes are not always guaranteed. (c.) Animals living in such an environment cannot always expect each foraging attempt to be successful. Thus, resources may not be accessible even though they are abundant. (d.) Expected resource availability, $\mathbb{E}(U)$ , is highest in areas with high resource abundance and high probability of success. When possible, animals are likely to rely mostly on predictable, resource-rich areas (top right). Alternatively, they may prefer predictable areas with lower resource abundance (top left) or adapt to high-risk, high-reward areas (bottom right). . . . .	11
6	Simulated spatial requirements for animals living in habitats where resource availability vary with constant, linearly increasing, cyclical, drifting, or erratic means and variances. The thick line indicates the mean spatial requirement needed for survival, while the thin line indicates the 95% quantile. The thick line can also be interpreted as the animal's core home range, while thin line can be interpreted as the requirements needed for a 95% chance of survival, which would be similar to the animals' 95% home range. Changes in $\mathbb{V}(R)$ have greater impacts when $\mathbb{E}(R)$ or $\mathbb{V}(R)$ are low. . . . .	13

7	(a.) Fictitious means of an arbitrary random variable (e.g., resource abundance, productivity, temperature, precipitation) in a two-dimensional space. The mean value varies within small-scale patches (left), between small-scale patches (center) and large scale-patches (right). Thus, the perceived means depend on the spatial scale at which the observer can detect differences.	16
8	(b.) Environmental heterogeneity detected at different degrees of precision. Wider regions have a wider range of means, which results in higher heterogeneity. Thus, each cell in the left panel represents the heterogeneity of the 100 pixels in the cell, while the cells in the central panel indicate the heterogeneity in the 10,000 cells in each of the larger cells, and the right panel indicates the heterogeneity in all 90,000 pixels. (c.) Fictitious, scale-dependent, perceived heterogeneity. The left panel is calculated as in row (b), while heterogeneity in the central and right panels are calculated on the means in the central and right panels in row (a), respectively. Perceived heterogeneity is lower than true heterogeneity because heterogeneity is lost when averaging values. . . . .	16
8	Simulations depicting the effects of resource availability and stochasticity on spatial needs. Animals moved from the circles to nearby tiles until satiated. The labels indicate how many steps animals took to reach satiety. Note the higher spatial needs of animals in more unpredictable or resource-scarce environments. Resources were generated using a gamma random variable parameterized by independent mean and variance paramters, which represented the resource abundance and environmental stochasticity, respectively, even though the two are likely correlated in nature. . . . .	21
9	Visual representation of the estimated timeline for my Master's project. . . . .	26

## Acknowledgements

I would like to thank Dr. Noonan for offering me the opportunity to pursue a Master's degree with such an interesting project with cutting-edge statistical methods for movement ecology. I have already learned a lot and I expect I will learn much more.

I would like to thank Dr. Ford and Dr. Pither for agreeing to be on my committee. Your feedback and mentorship will help me ensure this project is relevant and directly applicable to current issues regarding conservation and biogeography.

I would like to thank Sandra Fox for providing me with resources to ensure my perspective and work are inclusive and supportive of different ways of knowing, including Indigenous Knowledge, and that my attitude remains anti-colonial and anti-imperial.

## **Dedication**

To the Lands that have nourished me and given me a place to live, learn, and play.

# 1 Chapter 1: Literature review and quantitative predictions

Blackwell (2007):

- “spatiotemporal correlation” and “heterogeneity” instead of “patchiness”
- sometimes patches are actually just correlated areas rather than real patches

Hanya *et al.* (2006): the home range of Japanese macaques (*Macaca fuscata*) varied between seasons due to changes in food abundance and quality

Tucker *et al.* (2019): large birds fly longer distances in more homogeneous areas

Kouba *et al.* (2017): non-monotonic, nonlinear relationship between prey abundance and home range size

## 1.1 Home ranges as proxies for animals’ needs

The ability to move provides animals with the capacity to respond rapidly and continuously to heterogeneous and changing environments. Animals may move to search for resources (e.g., as food, nutrients, water, heat, a new den or nesting spot), a mate, or a new group. Movement also allows animals to escape predation or dangerous competition, and it allows them to defend resources and territory, too. Thus, we can estimate changes in animals’ movement and space use as a proxy for many needs, rather than measuring changes in the various individual needs (Nathan *et al.*, 2008).

In this context, the concept of a home range (HR) has a long history in ecological research as an indicator of the space an animal requires to satisfy its essential requirements during a period of time (Burt, 1943, see figure 1). These include both energetic needs and reproductive needs (which are not limited to finding a mate, since offspring require energy a safe location to develop in, too), but exploratory movement outside the habitual HR are generally excluded. However, for an animal to have what we may consider a HR, the animal must remain in a stable “home” area for long periods of time (Noonan *et al.*, 2019b). That is, the animal must be range-resident and the HR must have a stable centroid. Stable centroids may be concrete locations such as the dens or nests of central place foragers (*sensu* Orians, 1979), or they may be abstract points such as the center of an individual’s foraging grounds. Thus, animals with an unstable centroid would not be appropriate for HR analysis. In addition, while an individual’s HR may change over time (e.g., following a forest fire or a flood), it should remain stable during the period of observation. Ideally, properties of the HR (e.g., range size and structure) are representative of any new HR the animal will move to if the current one becomes inhospitable. For instance, an animal with a HR area of 1 km<sup>2</sup> would be expected to occupy a comparably sized area if it were displaced by a fire, assuming that the new habitat is sufficiently similar.

Similarly, the movement of the individual within the HR (e.g., range crossing time  $\tau_p$ , directional persistence  $\tau_v$ ) are also expected to be representative. Nomadic (e.g., Morato *et al.*, 2016; Nandintsetseg *et al.*, 2019) or migrating (e.g., Jonzén *et al.*, 2006; Abrahms *et al.*, 2019; Geremia *et al.*, 2019) animals are thus generally not considered range resident, and tracking periods should be long enough to provide a representative sample of the animal’s spatial needs.

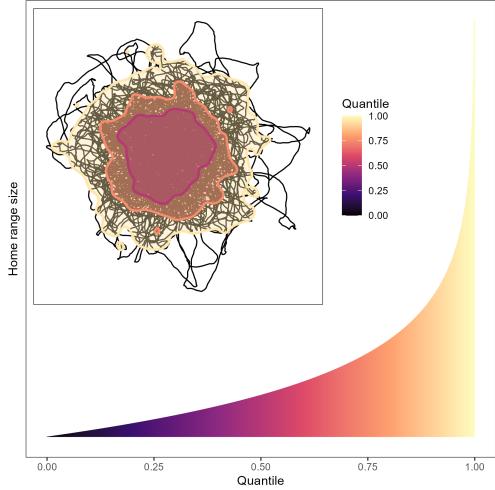


Figure 1: Inset: Simulated movement of an animal with a stationary home range centroid. The colored areas indicate the home range estimates for different quantiles. The red area indicates the core home range (50%), while the blue and green areas indicate the 75% and 95% home ranges, respectively. Main figure: Home range area as a function of utilization quantile. Note that as the quantile approaches 1, home range size approaches infinity, since if the animal was to move for an infinite amount of time, it would cover all possible space. In the case of a species or population where spatial needs change over the animals’ lifetimes (figure 2), one may define HR as a function of the animal’s life stages (with distinct discrete HRs) or as a function of age (so HR changes smoothly). In either case, the accuracy of the estimate will depend strongly on the length of the observation period as well as the measurement frequency (Noonan *et al.*, 2019b). Data from a portion of an animal’s life may be sufficient if it is representative the animal’s movement or if inference is limited to the period(s) for which data is available.

There are many factors which may affect animal’s spatial requirements and how they use their HR (*sensu* Nathan *et al.*, 2008). In particular, resource abundance is often assumed to be inversely proportional to HR size (or some function of it), such that regions with higher abundance correspond to smaller spatial needs, since animals do not have to range over extensive areas to meet the energetic requirements. However, the effect of many other factors likely depends on how an animal responds to them. For instance, competition may push individuals to explore other areas and expand their HR (Jetz *et al.*, 2004), but strong and consistent competition paired with territorial defense (e.g., wolves, Rich *et al.*, 2012; feral cats, Bengsen *et al.*, 2016;

capuchin monkeys, Tórrez-Herrera, Davis & Crofoot, 2020) may also prevent them from doing so. Similarly, predation may force animals to move more frequently to escape predators, or it may prevent them from venturing too far from the safety of their den (the HR's centroid) too often (Suraci *et al.*, 2022). Patch quality, size, fragmentation, and heterogeneity may cause animals to explore more patches if some are of low value, too small, too disconnected, or too variable (Fahrig *et al.*, 2019), but high diversity may also decrease HR size (Fox, 1981; Lucherini & Lovari, 1996). Similarly, patch connectivity and ease of movement may widen HRs by decreasing the energetic cost of movement and favoring exploration (Dickie *et al.*, 2022), or they may shrink HRs by decreasing the energetic cost of movement while increasing encounter rates with resources (Visser & Kiørboe, 2006; Bartumeus *et al.*, 2008; Martinez-Garcia *et al.*, 2020). However, not all animals take advantage of linear features or higher predictability in human-altered habitats. Noonan *et al.* (2021) found that giant anteaters (*Myrmecophaga tridactyla*) in Brazil did not use roads to reduce movement costs or increase movement speed. Instead, roads increased anteater mortality because the animals were attracted to the high-quality foraging found on roadside habitat. Similarly, Medici *et al.* (2022) found that the movement of tapirs (*Tapirus terrestris*) was unaffected by anthropogenic activity or habitat type.

The effects of resource abundance on animals' spatial needs have been studied by many in the last century. Multiple general hypotheses have been postulated based on (relatively) small-scale empirical studies (e.g., Burt, 1943; Southwood, 1977; Lindstedt & Boyce, 1985; Grant, 1993; Lucherini & Lovari, 1996; Nilsen, Herfindal & Linnell, 2005; Bengsen *et al.*, 2016), and supported (or questioned) by more recent work with larger, higher-resolution datasets (e.g., Jonzén *et al.*, 2006; Wolkovich *et al.*, 2012; Falcón-Cortés *et al.*, 2021; Dickie *et al.*, 2022; Nathan *et al.*, 2022) and simulations (Blackwell, 2007; Quaglietta, Porto & Ford, 2019; Tucker *et al.*, 2021). Recently, the amount of movement data which can be modeled at once has increased due to improvements in the quality and affordability of tracking equipment (Rutz & Hays, 2009), together with growing propensity (and requirements) to share data openly on various open data platforms such as Movebank (Kranstauber *et al.*, 2011; Kays *et al.*, 2022; but see Roche *et al.*, 2015), as well as the development

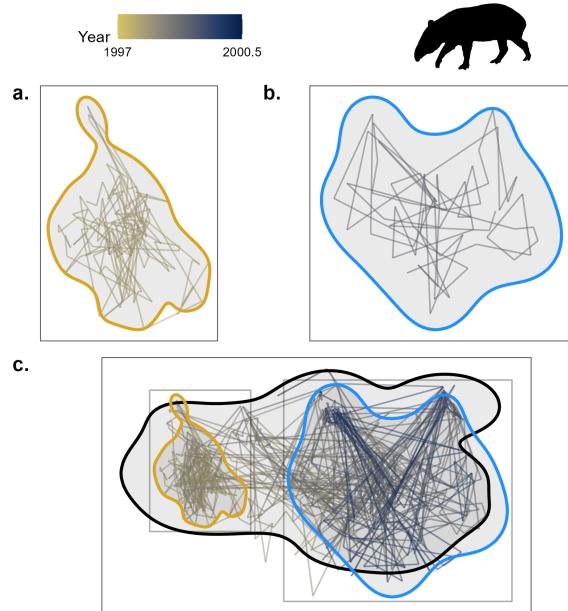


Figure 2: Movement of a tapir during a monitoring period of approximately 3 years (1997-07-10 to 2000-06-08) used in the work by Medici *et al.* (2022). Different subsamples may result in different estimates of home range size and centroid (a, b; each span 100 days), so neither subsample is necessarily representative of the space the animal used over the entire monitoring period (c).

of high-level modeling software (Bürkner, 2017, 2018; Wood, 2017; e.g., R Core Team, 2021). The abundance of data and statistical software allows researchers build on current knowledge by building increasingly complex hypotheses and models and testing them empirically and quantitatively.

While it is understood that an animal's spatial use strongly depends on the amount of resources and energy the animal can obtain from their habitat, estimates are often restricted to single populations or at most single species. To my knowledge, there are currently no large-scale estimations of vertebrate (or mammalian) space use as a function of resource availability. Additionally, little attention is often given to the stochasticity of resource availability or, more generally, habitat heterogeneity *and* stochasticity (but see Lucherini & Lovari, 1996; Nilsen *et al.*, 2005; Rizzuto *et al.*, 2021). This thesis aims to disentangle the effects of resource abundance and environmental stochasticity on animal space use using statistical models which are based on continuous-time stochastic processes and are insensitive to sampling frequency and spatiotemporal or taxonomic autocorrelation in the data. Findings from this thesis will provide information on how stochasticity has shaped the ecology and evolution of terrestrial mammals and how terrestrial mammals are currently adapting to heterogeneous and changing environments.

## 1.2 Effects of resource availability and productivity on spatial needs

Environmental productivity is tightly linked to the amount of space that animals need to cover to obtain the resources they needed to survive and reproduce (Lucherini & Lovari, 1996; Relyea, Lawrence & Demarais, 2000). While animals' needs vary greatly between taxonomic groups, some needs are essential for most species for survival and reproduction, such as energetic needs [e.g., food, water; see Baldwin & Bywater (1984)], habitat needs (e.g., dens, trees, tall grass, breeding grounds, protection from predators and competitors), and maintaining a thermoregulatory balance. The size of a home range, is hypothesized to be proportional to resource abundance (Burt, 1943), such that spatial needs increase when resources are low, but the relationship is likely not monotonic nor linear, since larger home ranges can result in higher rates of competition and are harder to defend (Grant, 1993; Jetz *et al.*, 2004).

The favorableness of a patch or habitat often depends on a variety of factors, including resource availability, competitive pressure, and predation risk. Let the random variable<sup>1</sup>  $R$  indicate the amount of *resources* in a particular patch, and let  $S$  be the random variable indicating whether or not a patch visit is *successful*. For simplicity, we can let  $S$  follow a Bernoulli distribution with probability of success  $p$  (which we can write as  $S \sim Ber(p)$ ). Next, let

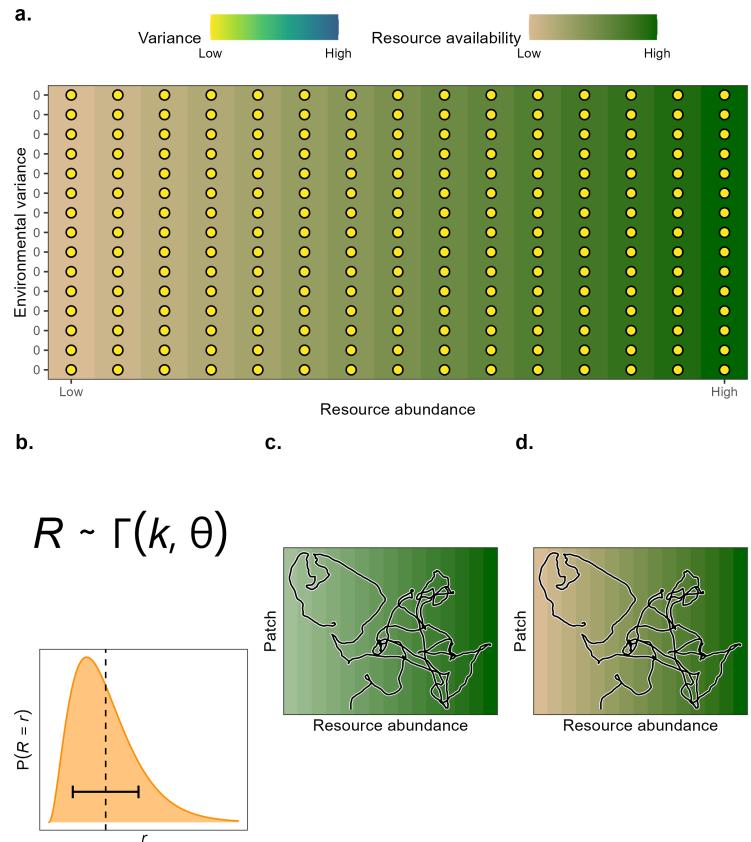


Figure 3: Fictitious example of variation in resource abundance in a heterogeneous but fully predictable environment. (a.) Complete predictability results in guaranteed successes (dots) during each foraging attempt. (b.) Arbitrary definition of  $R$  as following a Gamma distribution with shape  $k$  and scale  $\theta$ , while  $S$  follows a Bernoulli distribution with probability of success  $p = 1$ , since successes always occur (i.e.,  $P(S = 0) = 0$  implies  $S = 1$ ). Thus,  $U = R \cdot S = R$ , so  $U$  follows the same distribution as  $R$ . The shaded green area indicates the probability of obtaining an arbitrarily sufficient amount of resources to survive and reproduce. (c.) Since foraging attempts are guaranteed to be successful, the amount of usable resources will equal the resource abundance. (d.) Since  $U = R$ , the expected (i.e. average) amount of usable resources is equal to the expected resource abundance.

<sup>1</sup>In statistics, random variables indicate random (i.e., unknown) quantities and are indicated with capital letters (e.g.,  $R$ ,  $S$ ,  $U$ ). Known values, such as realizations (i.e., known observations) of random variables, are indicated with lower-case letters (e.g.,  $r$ ,  $s$ ,  $u$ ). Using this notation we can write the statement “the probability of random variable  $R$  taking the value  $r$ ” as  $P(R = r)$ .

$$U = R \cdot S \quad (1)$$

indicate the resources an animal can *use* during a visit. Following this simple model, a patch visit can result in two possible outcomes: if the visit is successful ( $S = 1$ ), an animal can use the entirety of the resource ( $U = R \cdot 1 = R$ ), but if it is unsuccessful ( $S = 0$ ) the animal is unable to access any of the resource ( $U = R \cdot 0 = 0$ ).

If we start by considering the simplistic, though admittedly unlikely, scenario where patches are fully predictable and free of disturbance and competition, such that the animal can access the patches' resources during any visit, i.e.,  $p = P(S = 1) = 1$ . In heterogeneous regions with no stochasticity (figure 3a-b), the favorableness of a location will depend strongly on  $R$ , so regions with higher  $R$  should be preferred (figure 3c). Since all patch visits in fully predictable regions are successful, animals can expect  $U$  to be equal to  $R$ , since  $U = R \cdot S = R \cdot 1 = R$  (figure 3d), which implies that  $\mathbb{E}(U) = \mathbb{E}(RS) = \mathbb{E}(R)$ . Since  $U = R$ , animals can choose their home ranges based on  $R$  directly, without having to account for any spatiotemporal stochasticity. Therefore, animals in such regions are likely to maximize fitness and minimize movement costs by spending large amounts of time in highly productive regions, with occasional exploratory movements to different patches. This model is quite simplistic, but its simplicity makes it easy to fit and conceptualize, and it provides a null model against which comparisons can be made. In other words, although it is unlikely for a region to be fully predictable, this model may provide acceptable results when environmental variance is low and has little effect on animal fitness, such as in areas where the variation in  $U$ ,  $\mathbb{V}(U)$ , is low relative to its expected value,  $\mathbb{E}(U)$ , such that the costs of moving to another favorable area are low or the chances of encountering prey are high. Mathematically, we can say that this simple model may be acceptable when the coefficient of variation,  $\sqrt{\mathbb{V}(R)} / \mathbb{E}(R)$ , is low. Additionally, this model may be the only option when data availability is too low to produce appreciable measures of stochasticity or there is no way to estimate it.

### 1.3 Effects of environmental variance on spatial needs

There are many sources of environmental change over both time and space. Some are due to repetitive, predictable, or even well-known patterns (such as daily or seasonal changes in temperature and precipitation or the location of different patches), while others are due to infrequent, unpredictable, or poorly-understood events (such as forest fires, the arrival of new competitors, or human activity). In this thesis, I will refer to predictable or known changes in space and time as environmental **heterogeneity** (defined as  $\mathbb{V}(R)$ ; see the orange distribution in figure 5), while I will use the term **stochasticity** specifically for unpredictable

(spatiotemporal) variation (defined as  $\mathbb{V}(S)$ ; see the blue distribution in figure 5)<sup>2</sup>. For example, the location of (high-yield) apple trees in an orchard and the time at which they produce fruit may be heterogeneous, but predictable. In contrast, the yield produced during a given year may be unpredictable (i.e., stochastic). Environmental sources of *heterogeneity* also include any known spatial variation in patch quality and size, and predictable daily or seasonal changes in temperature throughout the day, while environmental sources of *stochasticity*, include changes in temperature ranges, precipitation, and the frequency of extreme events due to climate change (IPCC, 2018; Noonan *et al.*, 2018). Unpredictable events such as forest fires, floods, and earthquakes also constitute sources of stochasticity.

Generally, events are predictable when (1) they occur with a probability density that is approximately constant over time and space, (2) they occur frequently within an organism's generation time or lifespan, and (3) they occur frequently enough to be expected as normal (e.g., rain in rainforests).<sup>3</sup> When an event occurs fairly frequently (e.g., approximately  $0.3 \lesssim p \lesssim 0.5$ ), animals may begin to expect the event to occur and consider it normality (figure 5a), as long as the change in frequency and magnitude is sufficiently gradual. For instance, Lamont *et al.* (2020) found that serotiny (the storage of seeds in closed cones or fruits which open following a fire) is common in plant populations which suffer fires at least once per lifespan, on average. However, the trait becomes less common if the fires become so common that plants survive more often as resprouters than by producing seeds. If a dangerous event such as fires becomes more frequent, unpredictable, or severe, organisms may store resources in favorable times and locations so they can resist more adverse times, and mobile animals may move to avoid such events altogether (Southwood, 1977; Lindstedt & Boyce, 1985). However, variety in fires may increase environmental heterogeneity and promote biodiversity (Fuhlendorf & Engle, 2004), particularly shortly after fires occur and when fires produce heterogeneous burns (Tingley *et al.*, 2016), but the effects of fire can differ between ecosystems (Koerner & Collins, 2013).

### 1.3.1 The temporal scale of stochastic events

The timescale over which an event is (un)predictable will also play a key role in determining how species respond, or not, to stochasticity. Adverse events which occur frequently during an organism's generation time or lifespan (e.g., at least twice, on average) may be considered as threats worth preparing for. Stochastic and destructive events which happen frequently can become part of an organism's habitat and life history, so organisms may either adapt to their occurrence and even learn to predict them (Foley, Pettorelli & Foley, 2008), or they may avoid the risk altogether by moving to a new habitat (Southwood, 1977). If habitats are

---

<sup>2</sup>Note that both sources of variation,  $\mathbb{V}(R)$  and  $\mathbb{V}(S)$  affect the variation in usable resource,  $\mathbb{V}(U)$ , which is represented in the green distribution in figure 5.

<sup>3</sup>The variance of a Bernoulli random variable  $Y$  is maximized when  $P(Y = 1) = p = 0.5 \implies \mathbb{V}(Y) = p(1 - p) = 0.25$ , and minimized when the event occurs almost never ( $p \approx 0$ ) or almost always ( $p \approx 1$ ), since now  $\mathbb{V}(Y) \approx 0(1 - 0) = 1(1 - 1) = 0$ . Thus, events which occur extremely rarely, such as meteor impacts, are predictably infrequent and often assumed not to occur.

(or become) unfavorable, animals may move to a new habitat, migrate seasonally (Geremia *et al.*, 2019), or simply prefer nomadism over range residency (e.g., Nandintsetseg *et al.*, 2019). It is unclear when animals switch from range residency to nomadism (or vice-versa), but understanding the connection between the two types of movement is important for quantifying the effect of spatiotemporal stochasticity on animal's spatial needs. From a quantitative perspective, the switch is related to an animal's positional autocorrelation over time (also known as home range crossing time, here indicated as  $\tau_p$ ). Animals without a constant home range (i.e., without a constant centroid) will spend more time away from the overall mean position, so the time required to crossing their entire range time will be large (i.e., on the order of the animal's lifespan).

Adverse events which are too infrequent (e.g.,  $p \lesssim 0.1$ ) may not be perceived as a reoccurring threat, so organisms may be unable to adapt appropriately. Instead, highly infrequent events are more likely to be perceived as an oddity rather than something worth preparing for. In contrast, highly stochastic events (e.g.,  $p \approx 0.5$  or if  $p$  changes unpredictably) may be perceived as a threat, but organisms may be unable to predict their occurrence, since it would require refined cognitive abilities. For a species or population to adapt to an event, the event must thus occur with sufficient frequency and for a sufficiently long period of time.

What an animal perceives as a stochastic process depends on the stochasticity of the process relative to the animal's size, current age (or average lifespan), and adaptability. Generally, small, short-lived, or young organisms will tend to be more sensitive to small-scale (spatial) or short-term (temporal) changes (Southwood, 1977). Smaller organisms (e.g., mice) are more likely to be severely impacted by a stochastic event than larger ones (e.g. elephants), since larger organisms can have bigger energy reserves (Lindstedt & Boyce, 1985), can move longer distances over short periods of time (Hirt *et al.*, 2017), and tend to have longer lifespans, generation times, and developmental periods (Brown *et al.*, 2004), which allow them to develop or memory about the frequency and severity of such events (Foley *et al.*, 2008; Polansky, Kilian & Wittemyer, 2015). However, the short generation time and high fertility of smaller *r-selected* species (Pianka, 1970; Brown *et al.*, 2004) can allow them to develop traits that increase the chances of survival following an extreme event. Additionally, the effects of size and lifespan on sensitivity are likely nonlinear and correlated, since smaller animals tend to have shorter lives (and vice-versa), and small animals also tend to have lower metabolic rates, which often limit their movement speed, home ranges, and how the animals interact with their ecosystem (Brown *et al.*, 2004). Thus, processes will have stronger impacts on smaller animals than on bigger ones: The grazing pattern of a bison (*Bison bison*) drastically alters the habitat of most crawling insects, but other grazing mammals would not pay attention to changes in grass length at the same spatial or temporal resolution. Similarly, the timing and quantity of yearly snowfall would be a somewhat predictable and expected event for most adult moose (*Alces alces*), but it may be a shock for many adult zooplankton

or a new-born wolf (*Canis lupus*). What one animal may perceive as a single stochastic event (or cycle) may be considered as a series of short and highly stochastic events by another animal.

Figure 4 illustrates the perceived resource availability mean and variance by animals with four different levels of adaptability and how they may predict the mean and variance to change in the future (after the vertical grey line). In the extreme (and likely fictitious) case where an animal does not adapt to the environment (dark blue), it cannot react to or predict changes in mean resource abundance, and thus the perceived variance is inflated whenever the resource abundance does not match the overall mean. Environments appear highly unpredictable to such animals. Animals which can only perceive simple, linear changes in mean resource abundance (orange) perform better as long as resources continue to change linearly in the same direction (i.e. increasing or decreasing). However, they may be surprised when the trend's direction changes, as indicated by the rapid increase in variance in the top rows of figure 4b. Animals which are able to adapt smoothly (light blue) can predict changes in mean without a process is not highly stochastic (unsurprisingly, see the bottom rows of figure 4). Finally, animals which adapt rapidly to changing environments may be most able to take advantage of frequent environmental changes, but they may be unable to produce reasonable predictions based on memory (*sensu* Fagan *et al.*, 2013; Abrahms *et al.*, 2019). Instead, they depend on constant information and only predict on the most recent information (as indicated by the deviations from the data when predicting into the future).

An animal's ability to alter its behavior (including its movement) in response to environmental conditions is essential in stochastic or changing landscapes. Whether the changes be due to highly variable but (potentially) predictable changes such as the seasons and the weather, or whether they be due to more stochastic events [e.g., natural events such as fires or floods, but also anthropogenic events like oil spills; see Matkin *et al.* (2008)], an animal's ability to adapt increases its odds of survival. Generally, temporal variation is more

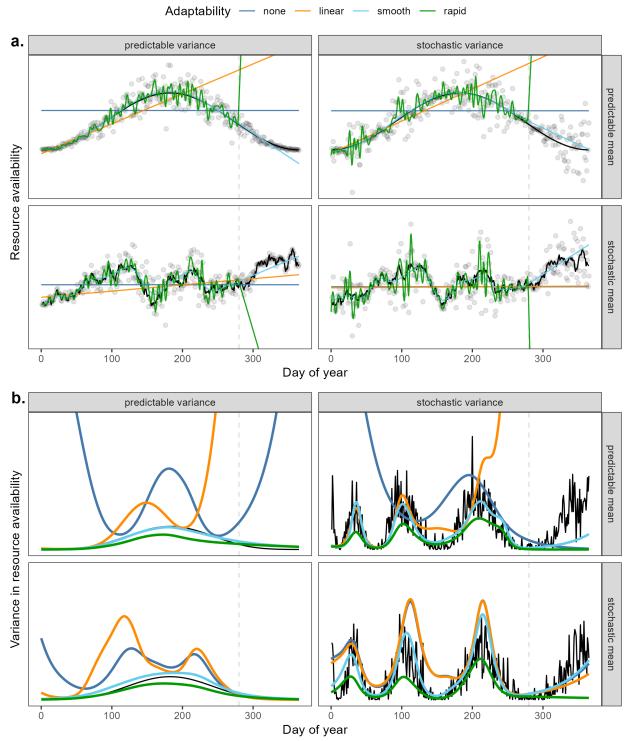


Figure 4: Fictitious changes in resource abundance mean (a) and variance (b) over the course of a year. The true trends are indicated by the black line, while the colored lines indicate the estimates perceived based on different levels of adaptability. The dashed vertical line indicates the hypothetical current date, such that any data to the right of the line is unknown to the animals, so estimates on the right of the dashed line are extrapolations. The data were simulated using Gaussian noise to ensure mean-variance independence and modeled using Generalized Additive Models with an identity link function to allow linear responses.

a significant bias in estimated variance, as long as the

Finally, animals which adapt rapidly to changing environments may be most able to take advantage of frequent environmental changes, but they may be unable to produce reasonable predictions based on memory (*sensu* Fagan *et al.*, 2013; Abrahms *et al.*, 2019). Instead, they depend on constant information and only predict on the most recent information (as indicated by the deviations from the data when predicting into the future).

likely to promote plasticity over diversity, since adaptability will likely offer better odds of survival than temporally static diversity in a population or species (Bell *et al.*, 1993). Rickbeil *et al.* (2019) showed that the yearly migration of elk (*Cervus canadensis*) depends on variable environmental events and cues such as available forage biomass, hunting pressure, snow fall, and snow melt. Birds have also shown to change their migration as the climate changes, including large-scale, trans-Saharan migrants (Jonzén *et al.*, 2006). And while genetic diversity and polymorphisms also increase the odds of survival for a species (Cavedon *et al.*, 2022), this project will focus strictly on animal behavior.

In an environment that changes over time, organisms which depend on mutable cycles such as changes in temperature, precipitation, and resource availability are more likely to respond to environmental changes than organisms which depend purely on deterministic cycles (e.g., photoperiod). While this thesis focuses on the movement and spatial use of animals, the adaptability (or lack thereof) of non-animal organisms is also crucial. The ability of most animals to rely on visual cues and move accordingly greatly increases their adaptability and plasticity, particularly for those animals that are able to move large distances over short periods of time (e.g., flying birds and large vertebrates). Although this project will focus on vertebrate animals, such changes are important to consider because the ability of a specialist or obligate symbiont to shift its home range or adapt will likely depend strongly on its associate's ability to move or adapt, too.

### 1.3.2 Spatial needs in stochastic environments

In areas where animals are not guaranteed that the resources they find during one visit will be there the next time (figure 5a), stochasticity will have an appreciable effect on the location's favourableness. Patches with low or high  $p$  will be most predictable, since successes can be expected to be very rare (if  $p \approx 0$ ) or very common (if  $p \approx 1$ ). In contrast, patches will be most stochastic when the probabilities of success and failure are approximately the same (i.e.,  $p \approx 1 - p \implies p \approx 0.5$ , see figure 5b). In stochastic habitats,  $\mathbb{E}(U)$  will depend on  $\mathbb{E}(S)$  as well as  $\mathbb{E}(R)$ , since  $S$  is no longer constant (figure 5c). Now, expected usable resources become

$$\mathbb{E}(U) = \mathbb{E}(R S), \tag{2}$$

or

$$\mathbb{E}(U) = \mathbb{E}(R) \mathbb{E}(S) \tag{3}$$

if  $R$  and  $S$  are independent. This model can be applied to all mobile animals, including herbivores, carnivores, and omnivores. In the case of herbivores,  $p$  may indicate the chance of finding good forage, which may be absent if regeneration times are long, if competitors have already exhausted the resource, or following a fire. For carnivores,  $p$  may indicate the chance of feeding on prey, which may depend on encountering and killing some first (Suraci *et al.*, 2022). However, since  $R$  is not limited to energetic resources,  $p$  may also indicate the chance of finding water or also a suitable location for a den or nest.

While resource availability is often considered in ecology (Burt, 1943; Southwood, 1977; Relyea *et al.*, 2000; Nilsen *et al.*, 2005; Williams-Guillen *et al.*, 2006; Rickbeil *et al.*, 2019), an environment's heterogeneity (i.e., diversity), stochasticity (i.e., unpredictability), and how the

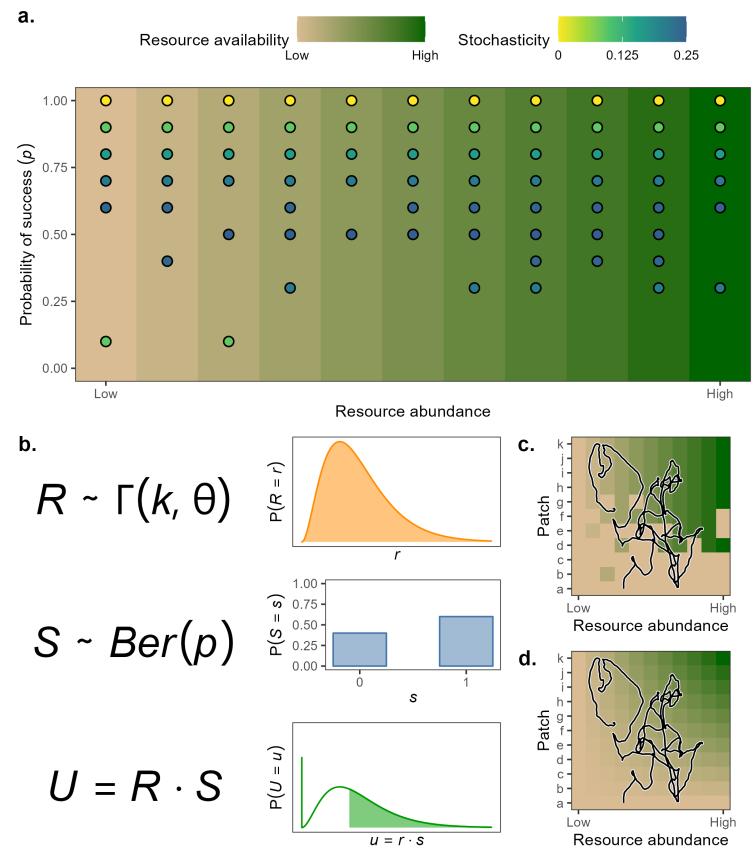


Figure 5: Fictitious example of variation in resource abundance in a heterogeneous and stochastic environment. (a.) Foraging successes (dots) can vary due to various reasons, including differences in competition, predation, and resource-specific trends. Successes most stochastic when the probability of success is approximately 0.5, while they are least stochastic (i.e., most predictable) when they are guaranteed to occur ( $p = 1$ ) or guaranteed to fail ( $p = 0$ ). (b.) Arbitrary definition of  $R$  as following a Gamma distribution with shape  $k$  and scale  $\theta$ , while  $S$  follows a Bernoulli distribution with probability of success  $p = 0.6$ , so successes are not always guaranteed. (c.) Animals living in such an environment cannot always expect each foraging attempt to be successful. Thus, resources may not be accessible even though they are abundant. (d.) Expected resource availability,  $\mathbb{E}(U)$ , is highest in areas with high resource abundance and high probability of success. When possible, animals are likely to rely mostly on predictable, resource-rich areas (top right). Alternatively, they may prefer predictable areas with lower resource abundance (top left) or adapt to high-risk, high-reward areas (bottom right).

two change over time are rarely accounted for. In addition, environmental stochasticity, including extreme events, can reduce a landscape's energetic balance (Chevin, Lande & Mace, 2010), which, in turn, decreases animals' fitness. Therefore, I expect animals living in unpredictable environments to require more space than those in stable environments. Although this hypothesis is supported by a few recent studies (Morellet *et al.*, 2013; Nandintsetseg *et al.*, 2019; Riotte-Lambert & Matthiopoulos, 2020), many of them are limited in their analytical depth and geographic and taxonomic scales, so the extent to which these preliminary findings can be generalized is still very limited. There thus remains a need for developing a more complete understanding of how animals' spatial needs change with environmental stochasticity. These stresses are compounded by climate change, which exposes species to increasingly common stochastic events (IPCC, 2018; Noonan *et*

*al.*, 2018). Furthermore, anthropogenic structures reduce the habitat available to terrestrial species (Wilson *et al.*, 2016), who struggle to move in fragmented (Fahrig, 2007), human-dominated landscapes (Tucker *et al.*, 2018). As the impacts of habitat loss and climate change will worsen in the future (Hansen *et al.*, 2013; IPCC, 2018), it is imperative that we better understand spatial requirements of taxa to protect wildlife existence and biodiversity. Environmental safeguarding is also essential for Reconciliation with Indigenous People in Canada (Truth and Reconciliation Commission of Canada, 2015).

## 1.4 Interaction effects of resource availability and stochasticity

In the previous section, I mentioned that a model which does not account for environmental variance may be acceptable when  $\sqrt{\mathbb{V}(R)}/\mathbb{E}(R)$  is low, as in the case of a habitat with high  $R$  and  $p \approx 1$ . However, since maximum resource abundance is constrained by environmental stochasticity (Chevin *et al.*, 2010) and resources can be depleted or rendered inaccessible by other individuals (Grant, 1993; Jetz *et al.*, 2004),  $\mathbb{E}(R)$  does not provide a sufficiently complete picture of resource availability in an environment. Thus, the effect of  $\mathbb{V}(R)$  should be included when possible, even when  $p \approx 1$  or  $p \approx 0$ .

Let  $H$  indicate the size of an animal's HR.<sup>4</sup> As explained above,  $H$  will be higher in regions with lower  $\mathbb{E}(U)$  or higher  $\mathbb{V}(U)$ . Figure 6 presents various scenarios with different trends in  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$ . Although some of these scenarios may seem overly simplistic, they can help us understand the complex interaction effects of  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$ . In regions where  $U$  has constant mean and variance,  $H$  remains approximately constant, although small oscillations are possible, since  $\mathbb{V}(U) \neq 0$  and thus the environment is still stochastic. When  $\mathbb{E}(U)$  changes over time but  $\mathbb{V}(U)$  remains constant, it is easy to see that  $H$  decreases as  $\mathbb{E}(U)$  increases, with changes in  $\mathbb{E}(U)$  having larger effects when  $\mathbb{E}(U)$  is low and smaller effects when  $\mathbb{E}(U)$  is already large. This is best visualized in the scenario in which  $\mathbb{E}(U)$  is increasing linearly while  $\mathbb{V}(U)$  is constant, since decrease in core HR ( $H_{50\%}$ ) and 95% HR ( $H_{95\%}$ ) is decreases over time. Additionally, as  $\mathbb{E}(U)$  increases, the difference between  $H_{50\%}$  and  $H_{95\%}$  also decreases. Thus,  $H_{95\%}$  is more sensitive to changes in  $\mathbb{E}(U)$  than  $H_{50\%}$ .  $H_{95\%}$  is also more sensitive to changes in  $\mathbb{V}(U)$ , since changes in  $\mathbb{V}(U)$  cause greater oscillations in  $H_{95\%}$  than in  $H_{50\%}$ .

Not all 25 scenarios depicted in figure 6 may be realistic, but the trends in  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$ , and their impacts on animal space use, are useful examples that can be thought of as simplified scenarios.  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$  may be (approximately) constant in highly homogeneous environments, or environments where resources are sufficiently available that changes in  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$  remain undetected. Although it is impossible for  $\mathbb{E}(U)$

---

<sup>4</sup>We can consider specific utilization quantiles of the HR, such as the core HR,  $H_{50\%}$ , or the 95% HR,  $H_{95\%}$ , but for simplicity I will refer to the entirety of the HR with  $H$ . Statistically, we can imagine  $H$  as having a probability distribution with support over the interval from zero (not included) to infinity (also not included), which we can indicate with the notation  $H \in (0, \infty)$ .

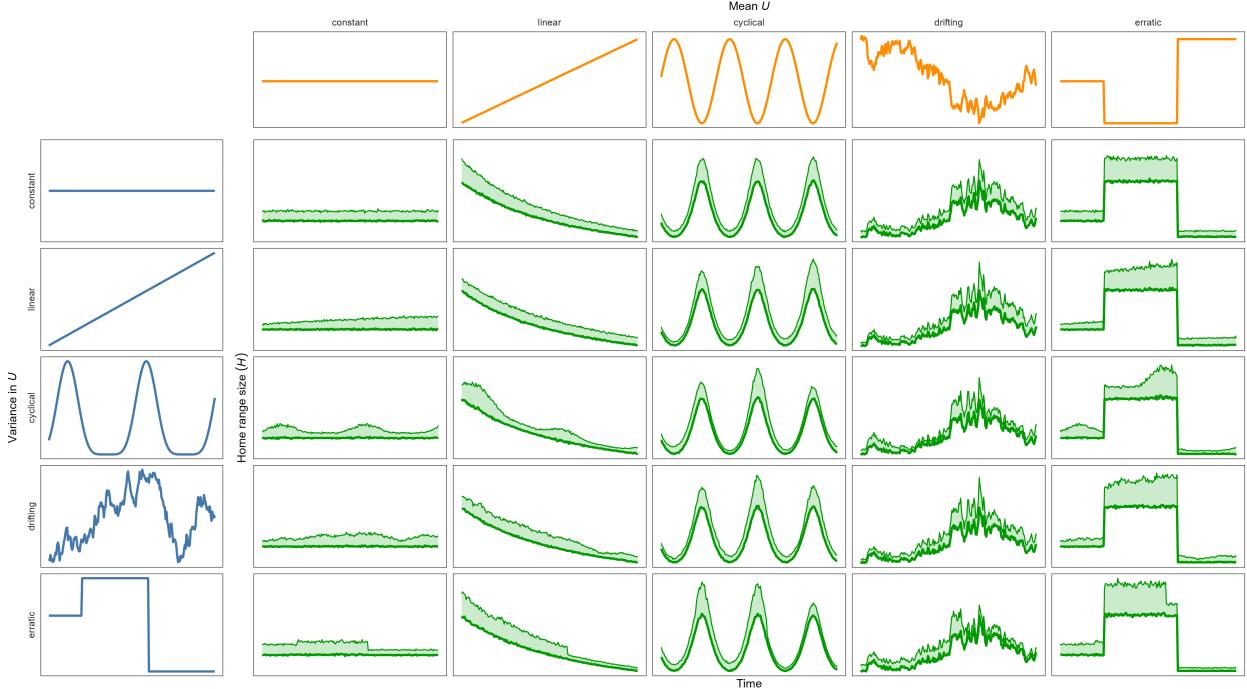


Figure 6: Simulated spatial requirements for animals living in habitats where resource availability vary with constant, linearly increasing, cyclical, drifting, or erratic means and variances. The thick line indicates the mean spatial requirement needed for survival, while the thin line indicates the 95% quantile. The thick line can also be interpreted as the animal's core home range, while thin line can be interpreted as the requirements needed for a 95% chance of survival, which would be similar to the animals' 95% home range. Changes in  $\mathbb{V}(R)$  have greater impacts when  $\mathbb{E}(R)$  or  $\mathbb{V}(R)$  are low.

and  $\mathbb{V}(U)$  to increase linearly continuously, such increases may be possible for short periods of time (followed by periods of no change or decrease). Additionally, these examples are important because they demonstrate the relationships between  $H$ ,  $\mathbb{E}(U)$ , and  $\mathbb{V}(U)$  in a (relatively) simple scenario. Cyclical oscillations in  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$  may occur in urban environments (Péron *et al.*, 2017) and as temperatures fluctuate daily and seasonally (Geremia *et al.*, 2019), while  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$  may drift randomly in highly complex environments with an abundance of competitors, threats, and stochasticity, such as a habitat with a high degree of human alteration and activity. Finally, erratic changes in  $\mathbb{E}(U)$  and  $\mathbb{V}(U)$  may occur in environments where changes are very sudden, such as fire-prone or flood-prone areas, or habitats with drastic human alteration (e.g., a forest which is clear-cut for mining purposes with a subsequent artificial re-forestation). However, if highly stochastic or erratic changes occur frequently, animals are most likely to perceive them as a smooth transition rather than a series of small, sudden, changes. Estimating the true trend would often require an excessively high cognitive capacity and an equally unlikely abundance of information. Additionally, although changes in  $\mathbb{E}(U)$  are not due to  $\mathbb{V}(U)$ , but often distinguishing between the two is not easy (Steixner-Kumar & Gläscher, 2020).

## 1.5 Thesis structure and aims

This thesis aims to quantify how animal's movement and use of space are affected by the abundance of resources (e.g., food, water, breeding grounds) and environmental stochasticity. This work has four key

objectives: (i) estimating individuals' spatial requirements in a way which is insensitive to variation in sampling protocols and data quality; (ii) quantifying environmental stochasticity and its effects; (iii) estimating between-species trends using models that are robust to commonly-found issues (e.g., correlations within species); and (iv) understanding how Traditional Indigenous Knowledge can be integrated into large-scale ecological research and conservation planning within a framework that acknowledges both Traditional Indigenous Knowledge and Western science (Kutz & Tomaselli, 2019). The present section provides the framework for this thesis, where each chapter is structured as a stand-alone body of work to be submitted for publication. While each chapter is designed to stand alone, together they provide convergent evidence towards the role of stochasticity in shaping animal space use.

To address these over-arching objections, in **Chapter 2**, I will produce a global raster of a new environmental stochasticity index, and new quantitative methods for animal movement. This metric will then serve as the basis for my subsequent investigation into the relationship between environmental stochasticity and animal movement.

In **Chapter 3**, I will use simulation studies and an unprecedented and conservation-relevant animal tracking dataset (>1500 animals, 77 globally-distributed species) to investigate how animal spatial needs change with environmental variance.

**Chapter 4** summarizes the work presented in this thesis and demonstrates its significance within the larger picture of movement ecology, conservation, and quantitative zoology.

Chapters 5-8 contain various supporting information, namely links to the code and data used in this thesis (**Chapter 5**), a table of all abbreviations used in the thesis (**Chapter 6**), a table of all mathematical notations and symbols used in the thesis (**Chapter 7**), and the tentative timeline (**Chapter 8**), which will not be present in the final version of the thesis.

**Appendix 1** contains details on further original work. These published works are included as appendix material as I have made substantial contributions, but I am not the lead author.

## 2 Chapter 2: A new measure of environmental variance

Notes:

- food quality may have a nonlinear relationship with NDVI, so resource use is also nonlinear over NDVI (Mueller *et al.*, 2008).  $\implies$  relationships should be modeled with models that do not assume linear trends but instead allow for smooth, nonlinear ones
- since NDVI accounts for photosynthetic activity, it contains info on phenology and seasonal cycles as well as the damage to vegetation caused by droughts, fires, floods, storms, blizzards, and frosts (Pettorelli *et al.*, 2011). Thus, we don't need to account for these extreme and rare events because their effects on vegetation are already accounted for (and are what likely matters most).
- NDVI as a proxy to estimate frequency of extreme events [e.g., fires, floods, droughts; see @]
- bison's grazing reduces fire risk and keeps nitrogen in the soil [[http://refhub.elsevier.com/S0169-5347\(19\)30052-7/sbref0480](http://refhub.elsevier.com/S0169-5347(19)30052-7/sbref0480)]

Whether an animal is affected by or can perceive environmental variance depends strongly on the spatiotemporal scale of the process(es) involved. Organisms are most affected by stochastic events and processes which occur on time scales which are in the order of the organisms' life spans or generation times (Southwood, 1977). Weekly heavy rains which alter a lake's salinity are more likely to affect an individual than a multi-centennial drought, and the high salinity that follows the drought may be perceived as the (stressful) standard by individuals which were born during or following the drought. In contrast, organisms can adapt to unpredictable heavy rains if they occur on a daily frequency. However, stochastic processes and events which occur on time scales that are longer than an organism's lifespan may still cause significant effects on a population's fitness and stability. Droughts which occur on the time scale of centuries or millennia (Haig *et al.*, 2013) are unlikely to affect organisms directly, but such events could still alter the population's habitat or breeding grounds enough to cause a population collapse or prevent individuals from reproducing in their habitual breeding region (or reproduce altogether).

Thus, the scale at which we quantify variance should depend on the spatiotemporal scale and sensitivity of the animals being studied. Since this project focuses on medium to large animals, I will focus on processes and events which occur on the temporal scale of hours to decades and spatial scale of meters to kilometers. I will assume small-scale variances such as heterogeneity in grass density at the centimeter scale or fluctuations in temperature over the span of a few minutes have little to no (measurable) effect on the spatial needs of the animals which I will be studying. Even if such small-scale variances had an effect, such an effect would be hard to detect or quantify because: (1) the effect size would be small (so large amounts of data would

be necessary for an appreciable result), and (2) the uncertainty in the tracking data limits the precision at which one is able to detect changes in movement in response to such small variation (figure 7). Thus, it seems reasonable to quantify environmental variance at the scale of minutes to days and meters to kilometers.

To my knowledge, there currently is no large-scale measurement of environmental variance. Producing a worldwide raster of such a measure would allow researchers to estimate the spatiotemporal variance of an animal’s habitat and better understand not only how heterogeneity and unpredictability affect its home range, but also when and why animals decide to migrate or become nomadic, and many other behaviors or decisions, such as the timing of reproduction, when animals decide to defend territory and resources, and more. Some measures of environmental productivity and heterogeneity already exist, such as the machine learning human footprint index (ml-HFI) produced by Keys, Barnes & Carter (2021) and the Normalized Difference Vegetation Index (NDVI, Pettorelli

*et al.*, 2011). However, none of these indices provides a comprehensive measure of environmental variance over the years. For instance, the ml-HFI is a temporally static raster, so it prevents us from accounting for temporal changes in anthropogenic habitat alteration, and neither the ml-HFI and the NDVI account for the frequency of events.

A measure of environmental variance should thus account for the spatiotemporal variance in frequent events (e.g., precipitation) and continuous processes (e.g., temperature) as well as the frequency of extreme and rare events (e.g., fires, cyclones). How the frequency and average intensity of events are estimated depends on the event(s) considered. Processes for which we have an abundance of global-scale data, such as daily air temperature and precipitation can be included as raw data, and their variances can be estimated using smooth

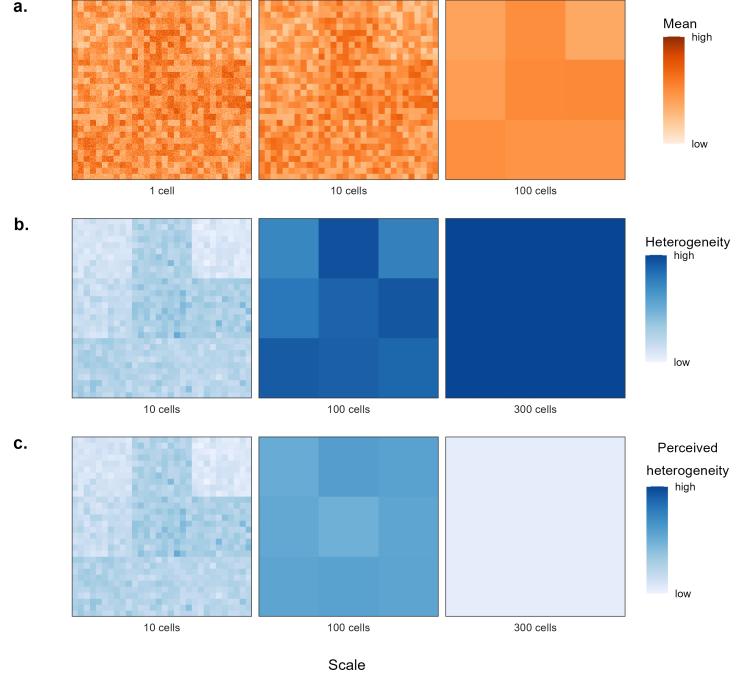


Figure 7: (a.) Fictitious means of an arbitrary random variable (e.g., resource abundance, productivity, temperature, precipitation) in a two-dimensional space. The mean value varies within small-scale patches (left), between small-scale patches (center) and large scale-patches (right). Thus, the perceived means depend on the spatial scale at which the observer can detect differences. (b.) Environmental heterogeneity detected at different degrees of precision. Wider regions have a wider range of means, which results in higher heterogeneity. Thus, each cell in the left panel represents the heterogeneity of the 100 pixels in the cell, while the cells in the central panel indicate the heterogeneity in the 10,000 cells in each of the larger cells, and the right panel indicates the heterogeneity in all 90,000 pixels. (c.) Fictitious, scale-dependent, perceived heterogeneity. The left panel is calculated as in row (b), while heterogeneity in the central and right panels are calculated on the means in the central and right panels in row (a), respectively. Perceived heterogeneity is lower than true heterogeneity because heterogeneity is lost when averaging values.

location-scale models, such as GAMs with appropriate conditional distributions [i.e., Gaussian location-scale for temperature and Gamma location-scale for precipitation; see Stasinopoulos & Rigby (2007); Umlauf, Klein & Zeileis (2018)]. In contrast, the frequency of rare events should be modeled with more smooth (i.e., less flexible) models, since the low data availability does not allow us to use models with high degrees of freedom [i.e., very wiggly predictors; see Simpson (2018)].

## 2.1 Quantifying environmental variance

Let  $Z$  indicate the measure of environmental variance used in this thesis.  $Z$  should contain information on a multitude of sources of environmental variance, including daily, seasonal, and spatial variance (e.g., changes in temperature, precipitation, and photoperiod, NDVI), inter-annual variance (e.g., climatic oscillations such as the El Niño-Southern Oscillation, see Rasmusson & Wallace, 1983; and anthropogenic climate change, see IPCC, 2018), and the frequency of extreme events, such as fires, floods, and droughts. The variation from most sources can be modeled using a Generalized Additive Model (GAM, see Wood, 2011; Wood, 2017) for location and scale (GAMLS, see Stasinopoulos & Rigby, 2007) with an appropriate distribution.<sup>5</sup> Similarly, the frequency of extreme events can be estimated from the estimated probability of occurrence of the event. We can estimate such probability using a piece-wise exponential additive model [PAM; see Bender & Scheipl (2018)], which is a GAM for time-to-event data.

Since each source of variation will affect animals to different degrees and will have different units (e.g., °C, milliliters, hours) or no units (e.g., NDVI, frequency of events), and different ranges (e.g.,  $(-\infty, \infty)$  for temperature,  $[0, \infty)$  for precipitation,  $[0, 1]$  for fraction of daylight, and  $[-1, 1]$  for NDVI), it would not be appropriate to define  $Z$  as the sum of each variance. Instead, a weighted average of the (scaled) variances or coefficients of variation (i.e., standard deviation over mean) may be more appropriate. Weights and scaling could be informed by an algorithm similar to a principal components analysis (PCA) that varies over time and space. However, the more abstract  $Z$  is, the less interpretable it becomes, so results based on  $Z$  should be interpreted carefully since a different set of variables or different weights may result in different conclusions. Therefore, although using few sources of variance will fail to account for some factors, the results may be easier to understand and apply. This would particularly be the case if few variables accounted for most environmental variance.

---

<sup>5</sup>For example, temperature (measured in °C) could be modeled using a Gaussian distribution since it can take both positive and negative values, while precipitation is strictly positive and continuous, so a Gamma distribution would be appropriate. A beta distribution could be used to model the proportion of a day with light (e.g., hours of sunlight divided by 24, which would result in a proportion in the interval  $[0, 1]$ ) and NDVI (Pettorelli *et al.*, 2011), since it can easily be transformed from the interval  $[-1, 1]$  to  $[0, 1]$  using the function  $g(x) = (x + 1)/2$ . Note that it is acceptable to transform the models' responses (light per day and NDVI) since both transformations ( $f(x) = x/24$  and  $g(x) = (x + 1)/2$ ) are linear (i.e. only comprised of addition and multiplication), so Jensen's inequality (Jensen, 1906) does not apply here.

Since the variance of a process has a lower limit (0) but does not necessarily have an upper limit, it seems reasonable to model  $Z$  using a Gamma distribution for the model residuals. Estimating  $\mathbb{E}(Z)$  over time and space will require various measures of spatiotemporal heterogeneity and stochasticity. The measure should vary over time and space, such that: (1)  $\mathbb{E}(Z)$  changes over time, (2) individual location can have different values of  $\mathbb{E}(Z)$ , and (3) the change in  $\mathbb{E}(Z)$  over time can also differ between habitats, since it may change faster in some habitats than others, and the shape and direction of the change may also differ between locations. Thus, the model used to estimate environmental variance should have 3 main terms: (1) a term for the global average over time ( $t$ ), (2) a term for the global average over space (longitude,  $s_x$ , and latitude,  $s_y$ ), and (3) an interaction term of space and time that allows locations to deviate from the global average temporal and spatial trends:

$$\mathbb{E}(Z) = f_1(t) + f_2(s_1, s_2) + f_3(t, s_1, s_2) \quad (4)$$

The terms should be smooth and sufficiently complex to maximize model flexibility, but the wigginess should also be penalized to avoiding over-fitting the data (Simpson, 2018). Thus, it seems appropriate to estimate the average  $Z$  over space and time using a GAM, which can be fit easily using the `mgcv` (Wood, 2017) package for R (R Core Team, 2021). The smooth of time,  $f_1(t)$  could be modeled with a thin-plate spline smooth (Wood, 2003; Simpson, 2018), while the two-dimensional spatial term,  $f_2(s_1, s_2)$ , should use a two-dimensional spline. While splines on the sphere (Wahba, 1981) would provide a good approximation for the shape of the earth,  $Z$  will likely be constrained to terrestrial or freshwater habitats, so Duchon splines are a better alternative (Duchon, 1977). Duchon splines are well-behaved as they move away from the support of the data, so they are less likely to produce questionable edge behavior. Finally, the interaction term  $f_3(t, s_1, s_2)$  can be modeled as the tensor interaction product (Wood, Scheipl & Faraway, 2013) of time and space, with a thin-plate spline smooth for time and a Duchon spline for space. To reduce computation time, the GAM will be fit using the `bam()` function from the `mgcv` package, which fits models similarly to `mgcv::gam()` but is specific for fitting models to very large datasets. `bam()` allows model fitting using fast restricted maximum likelihood and discretized covariates (Wood, Goude & Shaw, 2015; Wood *et al.*, 2017; Li & Wood, 2020), with no appreciable loss to model fit.

### 3 Chapter 3: Movement analyses

(Noonan *et al.*, 2019a) \* need  $\Delta t < \tau_v$ . If  $\Delta t > \tau_v$ , speed estimates become unreliable. e.g., estimating velocities from the movement of a pika that changes direction every 30 seconds will require more frequent sampling, while sampling the movement of a migrating goose that flies in a straight line for an hour or more (ref?) will require less frequent measurements. If  $\Delta t > 3\tau_v$ , since directional persistence decays exponentially at rate  $1/\tau_v$ , the time required for the proportion of the original velocity autocorrelation to decay to statistically insignificant levels (e.g.,  $\alpha > 0.05$ , but see Amrhein, Greenland & McShane, 2019) is  $\tau_\alpha = \tau_v \log(1/\alpha) \implies \tau_{0.05} = \tau_v \log(1/0.05) \approx 3$  \* the scale sensitivity of discrete movement models causes them to under-estimate speed if the sampling is too coarse and over-estimate speed if the sampling is too fine \* continuous-time estimates of speed and distance require proper error calibration, otherwise the models will produce very biased estimates (even worse than SLD estimates) \* SLD is ridiculously sensitive to data frequency \* good estimates of speed require  $\Delta t \ll 3\tau_v$  and a high  $N_{velocity}$ , which are the measurements with independent velocities (i.e.,  $t_i - t_{i-1} \gtrsim 3\tau_v$ )

- dingos that feed on human scraps have a smaller HR: <https://doi.org/10.1111/j.1600-0587.2013.00056.x>

This chapter will contain the analyses of the animal movement dataset which is central to this thesis, including how the animals' spatial needs and behaviors were quantified. It will then use the results and products from the previous chapters (namely the movement simulations and the raster of environmental variance) to estimate the effects of resource abundance and environmental variance on animals' spatial needs. Animal movement data will be modeled using continuous-time (stochastic) movement models (CTMMs), which assume movement data to be generated from stochastic process with spatiotemporal autocorrelation. In contrast, most models used in movement ecology (e.g., minimum convex polygon, Mohr, 1947; kernel density estimation, Worton, 1989; k-local convex hull, Getz & Wilmers, 2004) assume movement processes to be temporally discrete (e.g., an animal moves at the same time every day) and spatiotemporally independent (i.e., the current position of an animal now does not affect where it will be). The assumption of temporally discrete observations can be satisfied easily with a consistent data collection procedure (or data thinning), while the assumption of spatiotemporal independence may hold if the time between observations is particularly long (relative to the animal's  $\tau_p$ ). However, it is not hard to demonstrate that animal movement is not a discrete process, nor is it spatiotemporally independent: The position an animal is in at any given moment depends on where it was before, and it will affect where it will be later. Thus, when observations are taken with sufficient frequency, neither assumption is valid.

While the assumptions of discrete and independent observations may be acceptable (and even necessary) for

small datasets which do not contain sufficient information to estimate autocorrelation structures (DE Solla, Bonduriansky & Brooks, 1999; Blundell, Maier & Debevec, 2001; Fieberg, 2007), they are not appropriate for any high-frequency dataset, since such data is likely autocorrelated (Noonan *et al.*, 2019b). This is particularly the case with high-frequency data, since modern animal tracking datasets often have a lag-1 autocorrelation  $\approx 1$  (i.e., two consecutive observations are almost identical). In such cases, the sample size (i.e., the number of observations, often indicated as  $n$ ) of a dataset is not representative of the amount of information the dataset includes, so the concept of an *effective* sample size is preferred (indicated here as  $n_{eff}$ ). The effective sample size indicates the estimated number of independent observations, so it decreases as autocorrelation increases. When autocorrelation is not accounted for and the data is highly autocorrelated (which implies  $n_{eff}$  is much smaller than  $n$ ), any estimated measure of uncertainty is underestimated, so home range estimates tend to be artificially small (Noonan *et al.*, 2019b).

However, high autocorrelation is not an issue, as long as it is accounted for. Instead, it provides a more realistic representation of animal movement. Coarsening data until sampling occurs to regular intervals so autocorrelation is (arbitrarily) statistically not significant removes important properties of the dataset and nullifies great part of the benefits achieved with recent improvements in tracking technologies. Thus, it is increasingly important to model animal movement data using models that (1) do not assume data is regularly sampled, and (2) account for the spatiotemporal autocorrelation within the data.

Continuous-time (stochastic) movement models (CTMMs) such as Ornstein-Uhlenbeck (OU) and OU foraging (OUF) models relax the assumption of spatiotemporal independence by accounting for positional autocorrelation (OU and OUF models) and directional (i.e., velocity) autocorrelation (OUF models only). This allows the models to estimate the animal's average home range crossing time (OU and OUF) and the animal's average directional persistence (OUF only). Péron *et al.* (2017) provide additional information OU and OUF models and how to interpret them while also demonstrating how to use high-frequency movement data can help detect small-scale cycles, such as patterns that occur daily or weekly.

In this thesis, each animal's movement will be modeled using individual CTMMs, while the effects of resource abundance and environmental variance on various movement parameters (HR size, average speed,  $\tau_p$ ,  $\tau_v$ ) will be estimated using Hierarchical Generalized Additive Models (HGAMs, see Pedersen *et al.*, 2019). While it would be ideal to fit all movement data with a single hierarchical CTMM to use a single likelihood distribution and estimate common trends between animals and variances between animals' movement parameters, the `ctmm` package does not currently support hierarchical models. However, uncertainty in parameter estimates can still be accounted for in the HGAM by using some function of the parameters' standard error or  $n_{eff}$  as a weight (such that estimates with less certainty or independent data have less importance).

I expect home range size to decrease with resource abundance and increase with environmental variance, although environmental variance may have the most effect when  $R$  is low and a much lower effect when  $R$  is high (figure 8). Thus, the HGAM will include (1) a smooth predictor for resource abundance, (2) a smooth predictor of environmental variance, and (3) an smooth interaction effect of the two. The smooth interaction effect will allow the effect of environmental variance to depend on resource abundance (and vice-versa). Additionally, I plan to account for common trends between taxonomic groups by allowing each smooth effect to vary between taxonomic groups, but with a common degree of smoothness between groups (Pedersen *et al.*, 2019). If there is no evidence of bias arising from the opportunistic sampling of the dataset, I may also include spatiotemporal smooth terms to allow the effects of resource abundance and environmental variance to change over time and space. However, it may be more appropriate to fit such a model to a subset of the data to minimize sampling bias.

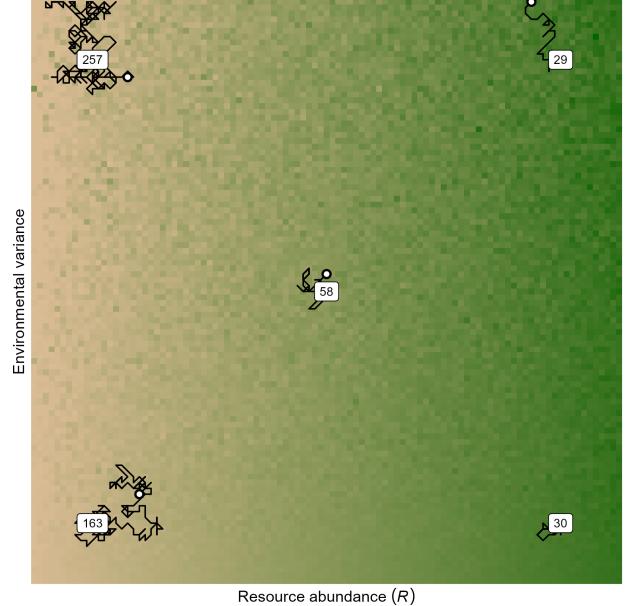


Figure 8: Simulations depicting the effects of resource availability and stochasticity on spatial needs. Animals moved from the circles to nearby tiles until satiated. The labels indicate how many steps animals took to reach satiety. Note the higher spatial needs of animals in more unpredictable or resource-scarce environments. Resources were generated using a gamma random variable parameterized by independent mean and variance parameters, which represented the resource abundance and environmental stochasticity, respectively, even though the two are likely correlated in nature.

## 4 Chapter 4: Synthesis

The effect of resource abundance on animals' spatial needs has been studied in multiple species, but little work has been done to estimate common trends between various taxonomic groups. Additionally, the spatiotemporal variability of the ecosystems animals live in is rarely taken into consideration. This thesis will provide estimates of how the spatial needs and behaviors of animals are affected by resource abundance and environmental variance.

**Chapter 1** will contain numerical simulations of how animals' movement and space use are affected by resource abundance and environmental variance, as well as a literature review that will help define the models used to produce such simulations. The review will also help determine which sources of environmental variance should be included in the global stochasticity raster presented in **Chapter 2**. Finally, in **Chapter 3**, I will use simulation studies and an unprecedented and conservation-relevant animal tracking dataset (>1500 animals, 77 globally-distributed species) to investigate how animals' spatial needs and behavior change with environmental variance.

The findings produced by this thesis will offer important considerations regarding the amount and quality of habitat animals require to survive in heterogeneous, changing, or stochastic habitats. Although the focus of this thesis will be on (large) mammals (and thus conclusions may not be applicable to other animals), the methods presented here can easily be applied to non-mammalian data. In addition, the global measure of environmental variance produced in **Chapter 2** will be designed so it is applicable to other fields of biology other than animal movement ecology, including animal behavior and phenology (including plant phenology).

## **5 Open and transparent science: Code and data availability**

All code and figures used in this project are available at the GitHub repository located at <https://github.com/StefanoMezzini/hr-environ-stoch-masters>. The repository currently does not include the animal movement data that will be used, which is available on Movebank and will be linked in future R scripts in the repository. The tapir movement data used for figure 2 is available in the `data` folder at <https://github.com/StefanoMezzini/tapirs>.

All scripts include comments to help people replicate the analysis, but they do assume some basic knowledge of R (including referring to help files and vignettes). Comments and requests regarding the project can be placed in the repository as issues.

## 6 List of abbreviations used

Table 1: A list of all abbreviations and acronyms used in the thesis.

Abbreviation	Phrase
CTMM	Continuous-time movement model
GAM	Generalized additive model
GAMLS	Generalized additive model for location and scale
HGAM	Hierarchical generalized additive model
ml-HFI	Machine learning human footprint index
NDVI	Normalized difference vegetation index
PAM	Piecewise exponential additive model
PCA	Principal components analysis

## 7 List of notations and symbols used

Table 2: A list of all mathematical notations and symbols used in the thesis.

Symbol	Meaning
$a < b$	$a$ less than $b$
$a > b$	$a$ greater than $b$
$a \approx b$	$a$ approximately equal to $b$
$a \geq b$	$a$ greater than or equal to $b$
$a \leq b$	$a$ less or equal to $b$
$a \lesssim b$	$a$ approximately less than $b$
$c \in (a, b]$	$c$ is in the interval from $a$ (not included) to $b$ (included)
$c \propto k$	$c$ is proportional to $k$
$\mathbb{E}(\cdot)$	Expectation, mean
$F$	Favorableness or value of a region
$H$	Home range, spatial needs, required area
$n$	sample size
$n_{eff}$	effective sample size
$P(A)$	Probability of event $A$ , e.g., $A$ = successful foraging
$R$	Resources
$R \sim Gamma(k, \theta)$	$R$ follows a Gamma distribution with parameters $k$ and $\theta$
$\tau_p$	Positional autocorrelation, mean reversion to HR center, range crossing time
$\tau_v$	Velocity autocorrelation, directional persistance
$\mathbb{V}(\cdot)$	Variance

## 8 Project timeline



Figure 9: Visual representation of the estimated timeline for my Master's project.

### 8.1 Progress to date

- Created and organized the GitHub repository (<https://github.com/StefanoMezzini/hr-environ-stoch-masters>);
- Completed all coursework necessary for a Master's degree;
- Reviewed literature and have begun making theoretical predictions;
- Living Data Project Working group: calculated new Canadian Living Planet Index (cLPI);
- Data analyst: estimated effects of human activity on tapir movement using `ctmm` and `mgcv`. The manuscript has been accepted and is in press (preprint is available at <https://www.biorxiv.org/content/10.1101/2021.11.12.468362v1>, and the code and data are available at <https://github.com/StefanoMezzini/tapirs>);
- Data analyst: estimated changes paleolimnological time series from endorheic lakes due to 8-m lake-level variation using Hierarchical Generalized Additive Models (HGAMs) and location-scale HGAMs fit with `mgcv`. The manuscript is available at <https://doi.org/10.1002/lno.12054>, while code and data are available at <https://github.com/simpson-lab/kenosee-white-bear>;
- Presented the background theory behind the thesis and some simulations as a guest lecturer in Dr. Noonan's BIOL 417 course (Evolutionary Ecology).

## Bibliography

- Abrahms B., Hazen E.L., Aikens E.O., Savoca M.S., Goldbogen J.A., Bograd S.J., *et al.* (2019). Memory and resource tracking drive blue whale migrations. *Proceedings of the National Academy of Sciences* **116**, 5582–5587. <https://doi.org/10.1073/pnas.1819031116>
- Amrhein V., Greenland S. & McShane B. (2019). Scientists rise up against statistical significance. *Nature* **567**, 305–307. <https://doi.org/10.1038/d41586-019-00857-9>
- Baldwin R. & Bywater A. (1984). Nutritional energetics of animals. *Annual review of nutrition* **4**, 101–114
- Bartumeus F., Catalan J., Viswanathan G.M., Raposo E.P. & Luz M.G.E. da (2008). The influence of turning angles on the success of non-oriented animal searches. *Journal of Theoretical Biology* **252**, 43–55. <https://doi.org/10.1016/j.jtbi.2008.01.009>
- Bélisle A.C., Asselin H., LeBlanc P. & Gauthier S. (2018). Local knowledge in ecological modeling. *Ecology and Society* **23**, art14. <https://doi.org/10.5751/ES-09949-230214>
- Bell G., Lechowicz M.J., Appenzeller A., Chandler M., DeBlois E., Jackson L., *et al.* (1993). The spatial structure of the physical environment. *Oecologia* **96**, 114–121. <https://doi.org/10.1007/BF00318038>
- Bender A. & Scheipl F. (2018). Pammtools: Piece-wise exponential Additive Mixed Modeling tools. *arXiv:1806.01042 [stat]*
- Bengsen A.J., Algar D., Ballard G., Buckmaster T., Comer S., Fleming P.J.S., *et al.* (2016). Feral cat home-range size varies predictably with landscape productivity and population density. *Journal of Zoology* **298**, 112–120. <https://doi.org/10.1111/jzo.12290>
- Bennett M.R., Bustos D., Pigati J.S., Springer K.B., Urban T.M., Holliday V.T., *et al.* (2021). Evidence of humans in North America during the Last Glacial Maximum. *Science* **373**, 1528–1531. <https://doi.org/10.1126/science.abg7586>
- Blackwell P.G. (2007). Heterogeneity, patchiness and correlation of resources. *Ecological Modelling* **207**, 349–355. <https://doi.org/10.1016/j.ecolmodel.2007.05.012>
- Blundell G.M., Maier J.A.K. & Debevec E.M. (2001). LINEAR HOME RANGES: EFFECTS OF SMOOTHING, SAMPLE SIZE, AND AUTOCORRELATION ON KERNEL ESTIMATES. *Ecological Monographs* **71**, 469–489. [https://doi.org/10.1890/0012-9615\(2001\)071%5B0469:LHREOS%5D2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071%5B0469:LHREOS%5D2.0.CO;2)
- Bolstad W.M. & Curran J.M. (2017). *Introduction to Bayesian statistics*, Third edition. Wiley, Hoboken, New Jersey.
- Bowles E., Marin K., MacLeod P. & Fraser D.J. (2021). A three-pronged approach that leans on Indigenous knowledge for northern fish monitoring and conservation. *Evolutionary Applications* **14**, 653–657. <https://doi.org/10.1111/eva.13146>

- Brown J.H., Gillooly J.F., Allen A.P., Savage V.M. & West G.B. (2004). TOWARD A METABOLIC THEORY OF ECOLOGY. *Ecology* **85**, 1771–1789. <https://doi.org/10.1890/03-9000>
- Bürkner P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal* **10**, 395. <https://doi.org/10.32614/RJ-2018-017>
- Bürkner P.-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80**. <https://doi.org/10.18637/jss.v080.i01>
- Burt W.H. (1943). Territoriality and Home Range Concepts as Applied to Mammals. *Journal of Mammalogy* **24**, 346. <https://doi.org/10.2307/1374834>
- Cavedon M., vonHoldt B., Hebblewhite M., Hegel T., Heppenheimer E., Hervieux D., et al. (2022). Genomic legacy of migration in endangered caribou. *PLOS Genetics* **18**, e1009974. <https://doi.org/10.1371/journal.pgen.1009974>
- Chevin L.-M., Lande R. & Mace G.M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLoS Biology* **8**, e1000357. <https://doi.org/10.1371/journal.pbio.1000357>
- DE Solla S.R., Bonduriansky R. & Brooks R.J. (1999). Eliminating autocorrelation reduces biological relevance of home range estimates. *Journal of Animal Ecology* **68**, 221–234. <https://doi.org/10.1046/j.1365-2656.1999.00279.x>
- Dickie M., Serrouya R., Avgar T., McLoughlin P., McNay R.S., DeMars C., et al. (2022). Resource exploitation efficiency collapses the home range of an apex predator. *Ecology*. <https://doi.org/10.1002/ecy.3642>
- Duchon J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: *Constructive Theory of Functions of Several Variables*. (Eds A. Dold, B. Eckmann, W. Schempp & K. Zeller), pp. 85–100. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Fagan W.F., Lewis M.A., Auger-Méthé M., Avgar T., Benhamou S., Breed G., et al. (2013). Spatial memory and animal movement. *Ecology Letters* **16**, 1316–1329. <https://doi.org/10.1111/ele.12165>
- Fahrig L. (2007). Non-optimal animal movement in human-altered landscapes. *Functional Ecology* **21**, 1003–1015. <https://doi.org/10.1111/j.1365-2435.2007.01326.x>
- Fahrig L., Arroyo-Rodríguez V., Bennett J.R., Boucher-Lalonde V., Cazetta E., Currie D.J., et al. (2019). Is habitat fragmentation bad for biodiversity? *Biological Conservation* **230**, 179–186. <https://doi.org/10.1016/j.biocon.2018.12.026>
- Falcón-Cortés A., Boyer D., Merrill E., Frair J.L. & Morales J.M. (2021). Hierarchical, Memory-Based Movement Models for Translocated Elk (*Cervus canadensis*). *Frontiers in Ecology and Evolution* **9**, 702925. <https://doi.org/10.3389/fevo.2021.702925>
- Fieberg J. (2007). KERNEL DENSITY ESTIMATORS OF HOME RANGE: SMOOTHING AND THE

- AUTOCORRELATION RED HERRING. *Ecology* **88**, 1059–1066. <https://doi.org/10.1890/06-0930>
- Foley C., Pettorelli N. & Foley L. (2008). Severe drought and calf survival in elephants. *Biology Letters* **4**, 541–544. <https://doi.org/10.1098/rsbl.2008.0370>
- Fox B.J. (1981). Niche Parameters and Species Richness. *Ecology* **62**, 1415–1425. <https://doi.org/10.2307/1941497>
- Fuhlendorf S.D. & Engle D.M. (2004). Application of the fire-grazing interaction to restore a shifting mosaic on tallgrass prairie: Shifting mosaic on tallgrass prairie. *Journal of Applied Ecology* **41**, 604–614. <https://doi.org/10.1111/j.0021-8901.2004.00937.x>
- Geremia C., Merkle J.A., Eacker D.R., Wallen R.L., White P.J., Hebblewhite M., et al. (2019). Migrating bison engineer the green wave. *Proceedings of the National Academy of Sciences* **116**, 25707–25713. <https://doi.org/10.1073/pnas.1913783116>
- Getz W.M. & Wilmers C.C. (2004). A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography* **27**, 489–505. <https://doi.org/10.1111/j.0906-7590.2004.03835.x>
- Girondot M. & Rizzo A. (2015). Bayesian Framework to Integrate Traditional Ecological Knowledge into Ecological Modeling: A Case Study. *Journal of Ethnobiology* **35**, 337–353. <https://doi.org/10.2993/etbi-35-02-337-353.1>
- Grant J.W.A. (1993). Whether or not to defend? The influence of resource distribution. *Marine Behaviour and Physiology* **23**, 137–153. <https://doi.org/10.1080/10236249309378862>
- Haig H.A., Kingsbury M.V., Laird K.R., Leavitt P.R., Laing R. & Cumming B.F. (2013). Assessment of drought over the past two millennia using near-shore sediment cores from a Canadian boreal lake. *Journal of Paleolimnology* **50**, 175–190. <https://doi.org/10.1007/s10933-013-9712-z>
- Hansen M.C., Potapov P.V., Moore R., Hancher M., Turubanova S.A., Tyukavina A., et al. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853. <https://doi.org/10.1126/science.1244693>
- Hanya G., Kiyono M., Yamada A., Suzuki K., Furukawa M., Yoshida Y., et al. (2006). Not only annual food abundance but also fallback food quality determines the Japanese macaque density: Evidence from seasonal variations in home range size. *Primates* **47**, 275–278. <https://doi.org/10.1007/s10329-005-0176-2>
- Hirt M.R., Jetz W., Rall B.C. & Brose U. (2017). A general scaling law reveals why the largest animals are not the fastest. *Nature Ecology & Evolution* **1**, 1116–1122. <https://doi.org/10.1038/s41559-017-0241-4>
- IPCC (2018). *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to*

*eradicate poverty.*

- Jensen J.L.W.V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica* **30**, 175–193. <https://doi.org/10.1007/BF02418571>
- Jetz W., Carbone C., Fulford J. & Brown J.H. (2004). The Scaling of Animal Space Use. *Science* **306**, 266–268. <https://doi.org/10.1126/science.1102138>
- Jonzén N., Lindén A., Ergon T., Knudsen E., Vik J.O., Rubolini D., *et al.* (2006). Rapid Advance of Spring Arrival Dates in Long-Distance Migratory Birds. *Science* **312**, 1959–1961. <https://doi.org/10.1126/science.1126119>
- Kays R., Davidson S.C., Berger M., Bohrer G., Fiedler W., Flack A., *et al.* (2022). The Movebank system for studying global animal movement and demography. *Methods in Ecology and Evolution* **13**, 419–431. <https://doi.org/10.1111/2041-210X.13767>
- Keys P.W., Barnes E.A. & Carter N.H. (2021). A machine-learning approach to human footprint index estimation with applications to sustainable development. *Environmental Research Letters* **16**, 044061. <https://doi.org/10.1088/1748-9326/abe00a>
- Kimmerer R.W. (2020). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge, and the teachings of plants*, Second hardcover edition. Milkweed Editions, Minneapolis.
- Koerner S.E. & Collins S.L. (2013). Small-scale patch structure in North American and South African grasslands responds differently to fire and grazing. *Landscape Ecology* **28**, 1293–1306. <https://doi.org/10.1007/s10980-013-9866-0>
- Kouba M., Bartoš L., Tomášek V., Popelková A., Šťastný K. & Zárybnická M. (2017). Home range size of Tengmalm's owl during breeding in Central Europe is determined by prey abundance. *PLOS ONE* **12**, e0177314. <https://doi.org/10.1371/journal.pone.0177314>
- Kranstauber B., Cameron A., Weinzerl R., Fountain T., Tilak S., Wikelski M., *et al.* (2011). The Movebank data model for animal tracking. *Environmental Modelling & Software* **26**, 834–835. <https://doi.org/10.1016/j.envsoft.2010.12.005>
- Kutz S. & Tomaselli M. (2019). “Two-eyed seeing” supports wildlife health. *Science* **364**, 1135–1137. <https://doi.org/10.1126/science.aaau6170>
- Lamont B.B., Pausas J.G., He T., Witkowski E.T.F. & Hanley M.E. (2020). Fire as a Selective Agent for both Serotiny and Nonserotiny Over Space and Time. *Critical Reviews in Plant Sciences* **39**, 140–172. <https://doi.org/10.1080/07352689.2020.1768465>
- Li Z. & Wood S.N. (2020). Faster model matrix crossproducts for large generalized linear models with discretized covariates. *Statistics and Computing* **30**, 19–25. <https://doi.org/10.1007/s11222-019-09864-2>
- Lindstedt S.L. & Boyce M.S. (1985). Seasonality, Fasting Endurance, and Body Size in Mammals. *The*

- American Naturalist* **125**, 873–878. <https://doi.org/10.1086/284385>
- Lucherini M. & Lovari S. (1996). Habitat richness affects home range size in the red fox *Vulpes vulpes*. *Behavioural Processes* **36**, 103–105. [https://doi.org/10.1016/0376-6357\(95\)00018-6](https://doi.org/10.1016/0376-6357(95)00018-6)
- Martinez-Garcia R., Fleming C.H., Seppelt R., Fagan W.F. & Calabrese J.M. (2020). How range residency and long-range perception change encounter rates. *Journal of Theoretical Biology* **498**, 110267. <https://doi.org/10.1016/j.jtbi.2020.110267>
- Matkin C., Saulitis E., Ellis G., Olesiuk P. & Rice S. (2008). Ongoing population-level impacts on killer whales *Orcinus orca* following the “Exxon Valdez” oil spill in Prince William Sound, Alaska. *Marine Ecology Progress Series* **356**, 269–281. <https://doi.org/https://doi.org/10.3354/meps07273>
- McElreath R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press/Taylor & Francis Group, Boca Raton.
- Medici E.P., Mezzini S., Fleming C.H., Calabrese J.M. & Noonan M.J. (2022). Movement ecology of vulnerable lowland tapirs between areas of varying human disturbance. *Movement Ecology* **10**, 14. <https://doi.org/10.1186/s40462-022-00313-w>
- Mohr C.O. (1947). Table of Equivalent Populations of North American Small Mammals. *American Midland Naturalist* **37**, 223. <https://doi.org/10.2307/2421652>
- Morato R.G., Stabach J.A., Fleming C.H., Calabrese J.M., De Paula R.C., Ferraz K.M.P.M., et al. (2016). Space Use and Movement of a Neotropical Top Predator: The Endangered Jaguar. *PLOS ONE* **11**, e0168176. <https://doi.org/10.1371/journal.pone.0168176>
- Morellet N., Bonenfant C., Börger L., Ossi F., Cagnacci F., Heurich M., et al. (2013). Seasonality, weather and climate affect home range size in roe deer across a wide latitudinal gradient within Europe. *Journal of Animal Ecology* **82**, 1326–1339. <https://doi.org/10.1111/1365-2656.12105>
- Mueller T., Olson K.A., Fuller T.K., Schaller G.B., Murray M.G. & Leimgruber P. (2008). In search of forage: Predicting dynamic habitats of Mongolian gazelles using satellite-based estimates of vegetation productivity. *Journal of Applied Ecology* **45**, 649–658. <https://doi.org/10.1111/j.1365-2664.2007.01371.x>
- Nandintsetseg D., Bracis C., Leimgruber P., Kaczensky P., Buuveibaatar B., Lkhagvasuren B., et al. (2019). Variability in nomadism: Environmental gradients modulate the movement behaviors of dryland ungulates. *Ecosphere* **10**. <https://doi.org/10.1002/ecs2.2924>
- Nathan R., Getz W.M., Revilla E., Holyoak M., Kadmon R., Saltz D., et al. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences* **105**, 19052–19059. <https://doi.org/10.1073/pnas.0800375105>
- Nathan R., Monk C.T., Arlinghaus R., Adam T., Alós J., Assaf M., et al. (2022). Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science* **375**, eabg1780. <https://doi.org/10.1126/science.eabg1780>

- //doi.org/10.1126/science.abg1780
- Nilsen E.B., Herfindal I. & Linnell J.D.C. (2005). Can intra-specific variation in carnivore home-range size be explained using remote-sensing estimates of environmental productivity? *Écoscience* **12**, 68–75. <https://doi.org/10.2980/i1195-6860-12-1-68.1>
- Noonan M.J., Ascensão F., Yogui D.R. & Desbiez A.L.J. (2021). Roads as ecological traps for giant anteaters. *Animal Conservation*, acv.12728. <https://doi.org/10.1111/acv.12728>
- Noonan M.J., Fleming C.H., Akre T.S., Drescher-Lehman J., Gurarie E., Harrison A.-L., *et al.* (2019a). Scale-insensitive estimation of speed and distance traveled from animal tracking data. *Movement Ecology* **7**, 35. <https://doi.org/10.1186/s40462-019-0177-1>
- Noonan M.J., Newman C., Markham A., Bilham K., Buesching C.D. & Macdonald D.W. (2018). In situ behavioral plasticity as compensation for weather variability: Implications for future climate change. *Climatic Change* **149**, 457–471. <https://doi.org/10.1007/s10584-018-2248-5>
- Noonan M.J., Tucker M.A., Fleming C.H., Akre T.S., Alberts S.C., Ali A.H., *et al.* (2019b). A comprehensive analysis of autocorrelation and bias in home range estimation. *Ecological Monographs* **89**, e01344. <https://doi.org/10.1002/ecm.1344>
- Orians G.H. (1979). On the theory of central place foraging. In: *Analysis of ecological systems*. pp. 157–177.
- Pedersen E.J., Miller D.L., Simpson G.L. & Ross N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **7**, e6876. <https://doi.org/10.7717/peerj.6876>
- Péron G., Fleming C.H., Paula R.C. de, Mitchell N., Strohbach M., Leimgruber P., *et al.* (2017). Periodic continuous-time movement models uncover behavioral changes of wild canids along anthropization gradients. *Ecological Monographs* **87**, 442–456. <https://doi.org/10.1002/ecm.1260>
- Pettorelli N., Ryan S., Mueller T., Bunnefeld N., Jedrzejewska B., Lima M., *et al.* (2011). The Normalized Difference Vegetation Index (NDVI): Unforeseen successes in animal ecology. *Climate Research* **46**, 15–27. <https://doi.org/10.3354/cr00936>
- Pianka E.R. (1970). On r- and K-Selection. *The American Naturalist* **104**, 592–597. <https://doi.org/10.1086/282697>
- Polansky L., Kilian W. & Wittemyer G. (2015). Elucidating the significance of spatial memory on movement decisions by African savannah elephants using state-space models. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20143042. <https://doi.org/10.1098/rspb.2014.3042>
- Quaglietta L., Porto M. & Ford A.T. (2019). Simulating animal movements to predict wildlife-vehicle collisions: Illustrating an application of the novel R package SiMRiv. *European Journal of Wildlife Research* **65**, 100. <https://doi.org/10.1007/s10344-019-1333-z>
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistics, Vienna, Austria.

- tical Computing, Vienna, Austria.
- Rasmusson E.M. & Wallace J.M. (1983). Meteorological Aspects of the El Niño/Southern Oscillation. *Science* **222**, 1195–1202. <https://doi.org/10.1126/science.222.4629.1195>
- Relyea R.A., Lawrence R.K. & Demarais S. (2000). Home Range of Desert Mule Deer: Testing the Body-Size and Habitat-Productivity Hypotheses. *The Journal of Wildlife Management* **64**, 146. <https://doi.org/10.2307/3802984>
- Rich L.N., Mitchell M.S., Gude J.A. & Sime C.A. (2012). Anthropogenic mortality, intraspecific competition, and prey availability influence territory sizes of wolves in Montana. *Journal of Mammalogy* **93**, 722–731. <https://doi.org/10.1644/11-MAMM-A-079.2>
- Rickbeil G.J.M., Merkle J.A., Anderson G., Atwood M.P., Beckmann J.P., Cole E.K., et al. (2019). Plasticity in elk migration timing is a response to changing environmental conditions. *Global Change Biology* **25**, 2368–2381. <https://doi.org/10.1111/gcb.14629>
- Riotte-Lambert L. & Matthiopoulos J. (2020). Environmental Predictability as a Cause and Consequence of Animal Movement. *Trends in Ecology & Evolution* **35**, 163–174. <https://doi.org/10.1016/j.tree.2019.09.009>
- Rizzuto M., Leroux S.J., Vander Wal E., Richmond I.C., Heckford T.R., Balluffi-Fry J., et al. (2021). Forage stoichiometry predicts the home range size of a small terrestrial herbivore. *Oecologia* **197**, 327–338. <https://doi.org/10.1007/s00442-021-04965-0>
- Roche D.G., Kruuk L.E.B., Lanfear R. & Binning S.A. (2015). Public Data Archiving in Ecology and Evolution: How Well Are We Doing? *PLOS Biology* **13**, e1002295. <https://doi.org/10.1371/journal.pbio.1002295>
- Rutz C. & Hays G.C. (2009). New frontiers in biologging science. *Biology Letters* **5**, 289–292. <https://doi.org/10.1098/rsbl.2009.0089>
- Schuster R., Germain R.R., Bennett J.R., Reo N.J. & Arcese P. (2019). Vertebrate biodiversity on indigenous-managed lands in Australia, Brazil, and Canada equals that in protected areas. *Environmental Science & Policy* **101**, 1–6. <https://doi.org/10.1016/j.envsci.2019.07.002>
- Simpson G.L. (2018). Modelling Palaeoecological Time Series Using Generalised Additive Models. *Frontiers in Ecology and Evolution* **6**, 149. <https://doi.org/10.3389/fevo.2018.00149>
- Smith L.T. (2021). *Decolonizing methodologies: Research and indigenous peoples*, Third edition. Zed, London.
- Southwood T.R.E. (1977). Habitat, the Templet for Ecological Strategies? *The Journal of Animal Ecology* **46**, 336. <https://doi.org/10.2307/3817>
- Stasinopoulos M.D. & Rigby R.A. (2007). Generalized additive models for location scale and shape

(GAMLSS) in R. *Journal of Statistical Software* **23**

Steixner-Kumar S. & Gläscher J. (2020). Strategies for navigating a dynamic world. *Science* **369**, 1056–1057.

<https://doi.org/10.1126/science.abd7258>

Suraci J.P., Smith J.A., Chamaillé-Jammes S., Gaynor K.M., Jones M., Luttbeg B., *et al.* (2022). Beyond spatial overlap: Harnessing new technologies to resolve the complexities of predator-prey interactions. *Oikos*. <https://doi.org/10.1111/oik.09004>

Tingley M.W., Ruiz-Gutiérrez V., Wilkerson R.L., Howell C.A. & Siegel R.B. (2016). Pyrodiversity promotes avian diversity over the decade following forest fire. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20161703. <https://doi.org/10.1098/rspb.2016.1703>

Tórrez-Herrera L.L., Davis G.H. & Crofoot M.C. (2020). Do Monkeys Avoid Areas of Home Range Overlap Because They Are Dangerous? A Test of the Risk Hypothesis in White-Faced Capuchin Monkeys (*Cebus capucinus*). *International Journal of Primatology* **41**, 246–264. <https://doi.org/10.1007/s10764-019-00110-0>

Truth and Reconciliation Commission of Canada (2015). *Honouring the truth, reconciling for the future: Summary of the final report of the Truth and Reconciliation Commission of Canada*. Truth; Reconciliation Commission of Canada.

Tucker M.A., Alexandrou O., Bierregaard R.O., Bildstein K.L., Böhning-Gaese K., Bracis C., *et al.* (2019). Large birds travel farther in homogeneous environments. *Global Ecology and Biogeography* **28**, 576–587. <https://doi.org/10.1111/geb.12875>

Tucker M.A., Böhning-Gaese K., Fagan W.F., Fryxell J.M., Van Moorter B., Alberts S.C., *et al.* (2018). Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science* **359**, 466–469. <https://doi.org/10.1126/science.aam9712>

Tucker M.A., Busana M., Huijbregts M.A.J. & Ford A.T. (2021). Human-induced reduction in mammalian movements impacts seed dispersal in the tropics. *Ecography* **44**, 897–906. <https://doi.org/10.1111/ecog.05210>

Umlauf N., Klein N. & Zeileis A. (2018). BAMLSS: Bayesian Additive Models for Location, Scale, and Shape (and Beyond). *Journal of Computational and Graphical Statistics* **27**, 612–627. <https://doi.org/10.1080/10618600.2017.1407325>

Visser A.W. & Kiørboe T. (2006). Plankton motility patterns and encounter rates. *Oecologia* **148**, 538–546. <https://doi.org/10.1007/s00442-006-0385-4>

Wahba G. (1981). Spline Interpolation and Smoothing on the Sphere. *SIAM Journal on Scientific and Statistical Computing* **2**, 5–16. <https://doi.org/10.1137/0902002>

Williams-Guillen K., McCann C., Martinez Sanchez J.C. & Koontz F. (2006). Resource availability and

- habitat use by mantled howling monkeys in a Nicaraguan coffee plantation: Can agroforests serve as core habitat for a forest mammal? *Animal Conservation* **9**, 331–338. <https://doi.org/10.1111/j.1469-1795.2006.00042.x>
- Wilson M.C., Chen X.-Y., Corlett R.T., Didham R.K., Ding P., Holt R.D., *et al.* (2016). Habitat fragmentation and biodiversity conservation: Key findings and future challenges. *Landscape Ecology* **31**, 219–227. <https://doi.org/10.1007/s10980-015-0312-3>
- Wolkovich E.M., Cook B.I., Allen J.M., Crimmins T.M., Betancourt J.L., Travers S.E., *et al.* (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature* **485**, 494–497. <https://doi.org/10.1038/nature11014>
- Wood S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semi-parametric generalized linear models. *Journal of the Royal Statistical Society (B)* **73**, 3–36
- Wood S.N. (2017). *Generalized additive models: An introduction with R*, Second edition. CRC Press/Taylor & Francis Group, Boca Raton.
- Wood S.N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**, 95–114. <https://doi.org/10.1111/1467-9868.00374>
- Wood S.N., Goude Y. & Shaw S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **64**, 139–155. <https://doi.org/10.1111/rssc.12068>
- Wood S.N., Li Z., Shaddick G. & Augustin N.H. (2017). Generalized Additive Models for Gigadata: Modeling the U.K. Black Smoke Network Daily Data. *Journal of the American Statistical Association* **112**, 1199–1210. <https://doi.org/10.1080/01621459.2016.1195744>
- Wood S.N., Scheipl F. & Faraway J.J. (2013). Straightforward intermediate rank tensor product smoothing in mixed models. *Statistics and Computing* **23**, 341–360. <https://doi.org/10.1007/s11222-012-9314-z>
- Worton B.J. (1989). Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology* **70**, 164–168. <https://doi.org/10.2307/1938423>

## **Appendix 1: Manuscripts contributed to during the degree**

## Abrupt changes in the physical and biological structure of endorheic upland lakes due to 8-m lake-level variation during the 20<sup>th</sup> century

Judith A. Bjorndahl,<sup>1,a</sup> Cale A.C. Gushulak ,<sup>1,2\*</sup> Stefano Mezzini ,<sup>2,b</sup> Gavin L. Simpson ,<sup>2,c</sup>  
 Heather A. Haig ,<sup>1,d</sup> Peter R. Leavitt ,<sup>1,2</sup> Kerri Finlay ,<sup>1</sup>

<sup>1</sup>Department of Biology, University of Regina, Regina, Saskatchewan, Canada

<sup>2</sup>Irving K. Barber School of Arts and Sciences, University of British Columbia, Kelowna, British Columbia, Canada

### Abstract

Climate-induced variation in lake level can affect physicochemical properties of endorheic lakes, but its consequences for phototrophic production and regime shifts are not well understood. Here, we quantified changes in the abundance and community composition of phototrophs in Kenosee and White Bear lakes, two endorheic basins in the parkland Moose Mountain uplands of southeastern Saskatchewan, Canada, which have experienced > 8 m declines in water level since ~ 1900. We hypothesized that lower water levels and warmer temperatures should manifest as increased abundance of phytoplankton, particularly cyanobacteria, and possibly trigger a regime shift to turbid conditions due to evaporative concentration of nutrients and solutes. High-resolution analysis of sedimentary pigments revealed an increase in total phototrophic abundance (as β-carotene) concurrent with lake-level decline beginning ~ 1930, but demonstrated little directional change in cyanobacteria. Instead, significant increases in obligately anaerobic purple sulfur bacteria (as okenone) occurred in both lakes during ~ 1930–1950, coeval with alterations to light environments and declines in lake level. The presence of okenone suggests that climate-induced increases in solute concentrations may have favored the formation of novel bacterial habitats where photic and anoxic zones overlapped. Generalized additive models showed that establishment of this unique habitat was likely preceded by increased temporal variance of sulfur bacteria, but not phytoplankton or cyanobacteria, suggesting that this abrupt change to physical lake structure was unique to deep-water environments. Such climate-induced shifts may become more frequent in the region due to hydrological stress on lake levels due to warming temperatures across the Northern Great Plains.

Local and regional declines in lake levels are of great concern as human water use is expected to increase over the next century (Vörösmarty et al. 2000; Gaeta et al. 2014). In

addition to anthropogenic uses, lake levels may vary in response to perturbations in regional hydroclimate, particularly where evaporation rates exceed precipitation levels (Pham et al. 2009; Xiao et al. 2018). Such hydroclimate changes and subsequent lake-level declines are widespread in the Northern Great Plains, particularly in hydrologically managed endorheic basins which depend on spring snowmelt for water replenishment (van der Kamp et al. 2008; Pham et al. 2009; Sereda et al. 2011). In these regions, general circulation models (GCMs) forecast warmer temperatures and only modest changes in precipitation (Tanzeeba and Gan 2012; Asong et al. 2016; Zhou et al. 2018), potentially resulting in increased volatility of regional lake levels due to evaporative forcing. Resultant changes in regional water quality may include variation in nutrient, carbon, and major ion concentrations, proportion of littoral and pelagic habitats, phytoplankton community composition and production, and food-web dynamics including zooplankton and fish communities (Vinebrooke et al. 1998; Fischer and Öhl 2005; Hambright et al. 2008; Pham et al. 2009; Gal et al. 2013; Wigdahl

\*Correspondence: cale.gushulak@uregina.ca, calegushulak@gmail.com

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Additional Supporting Information may be found in the online version of this article.

<sup>a</sup>Present address: Institute of Environmental Change and Society, University of Regina, Regina, Saskatchewan, Canada

<sup>b</sup>Present address: Department of Animal Science, Aarhus University, Aarhus, Denmark

<sup>c</sup>Present address: Saskatchewan Water Security Agency, Regina, Saskatchewan, Canada

<sup>d</sup>Present address: Department of Biology, Acadia University, Wolfville, Nova Scotia, Canada

Author Contribution Statement: J.A.B. and C.A.C.G. are equal lead authors.

et al. 2014; Vogt et al. 2018). Despite on-going variation in lake levels in the Canadian Prairie region (Fritz 1990; van der Kamp et al. 2008), little is known of the long-term effects of enhanced evaporative stresses on production, community composition, and stability of phototrophic assemblages.

Variation in lake levels also affects water-column heating and circulation, as well as the degree to which irradiance penetrates to stable environments such as lake bottoms or chemoclines (Zohary and Ostrovsky 2011). Changes in thermal stratification may be particularly noteworthy in shallow prairie and parkland systems where polymixis is common due to frequent high winds (Plancq et al. 2018), while variation in ionic content can affect chemical stratification (Hodgson et al. 1998). In addition, evaporative concentration of nutrients associated with lake-level decline may favor increased primary production, particularly under warm conditions (Vinebrooke et al. 1998; Zinabu 2002). As shallower ecosystems warm more quickly than deeper ones (Dröscher et al. 2009; Dibike et al. 2016), such conditions can promote blooms of planktonic algae and potentially-toxic cyanobacteria (Davis et al. 2009; Paerl and Paul 2012; Huisman et al. 2018; Hayes et al. 2019).

Large and disproportionate increases in cyanobacterial abundance can arise as an abrupt change, or “regime shift” (*sensu lato*) in some lake systems (Taranu et al. 2015; Bunting et al. 2016; Vogt et al. 2018). In some cases, there is increased temporal variance in the abundance or concentration of phytoplankton or cyanobacteria prior to the shift, marking a “critical slowing down” arising from changes in the strength of internal feedbacks due to environmental driver(s) (Scheffer et al. 2001; Scheffer and Carpenter 2003; Carpenter and Brock 2006; Dakos et al. 2015; Ratajczak et al. 2018). In lake systems, variation in primary production can rise because of prolonged gradual increases in nutrient influx (e.g., paradox of enrichment; Cottingham et al. 2000), after which the regime shift to prolific cyanobacteria is initiated by relative minor forcing that push lakes beyond critical thresholds (Dakos et al. 2015; Bunting et al. 2016). These transitions may become self-enforcing through positive feedback mechanisms and result in a transition to an alternate stable state that exhibits self-maintenance and hysteresis (Scheffer et al. 2001; Scheffer and Carpenter 2003; Dakos et al. 2015; Ratajczak et al. 2018). Additional environmental drivers which may result in a regime shift in lakes include changes in ice-cover duration, vertical-mixing regime, and water-column warming (Paerl and Paul 2012; Taranu et al. 2015); although, in theory, variation in lake level, chemical stratification, or ionic concentration could also induce a regime shift (Garcés et al. 1995; Hodgson et al. 1998). On the Northern Great Plains, changes in the interactions between the predominant air masses (Gulf, Pacific, Arctic) can cause increased evaporative concentration and salinity (Liu et al. 2008; Pham et al. 2009), variation in water-column mixing (polymixis to meromixis; Michels et al. 2007), and large changes (5–10 m) in lake level (van der

Kamp et al. 2008) that could all initiate a regime shift with large biological responses (Scheffer et al. 2001; Carpenter et al. 2011).

To date, little is known of whether changes in the physical status of lakes (deep/shallow, mixed/stratified, etc.) are recorded by temporal variation in primary production or whether such changes in physical conditions can initiate true regime shifts (Bunting et al. 2016; Taranu et al. 2018). Furthermore, it has been established that some regime shifts are not preceded by rising variance (Burthe et al. 2016; Ratajczak et al. 2018) and that rising variance in environmentally sensitive proxies does not invariably lead to a regime shift (Burthe et al. 2016), which makes the establishment of regime shifts very difficult without abundant a priori knowledge of the impacted system (Scheffer and Carpenter 2003; Dakos et al. 2015; Burthe et al. 2016). In this context, it may be useful to retroactively examine systems which have experienced substantial environmental changes to discover if any regime shifts occurred, either with or without an associated rise in variance of key proxies, such as biological production (Randsalu-Wendrup et al. 2016; Taranu et al. 2018). The establishment of past regime shifts, or transitions between alternate stable states, may provide critical insights on how nonlinear and abrupt changes in ecosystem response to environmental change may improve lake management and mitigation strategies (Bunting et al. 2016; Randsalu-Wendrup et al. 2016).

To better understand how lake-level variation may influence the temporal variability of lakes, we quantified historical changes in the production, composition, and variance of phototrophic assemblages in two proximal endorheic parkland lakes located in the Moose Mountain uplands of southeast Saskatchewan, Canada. Kenosee and White Bear lakes are presently unstratified (Plancq et al. 2018) and have experienced > 8 m declines in water level over the past century (Vance et al. 1997; Vinebrooke et al. 1998; van der Kamp et al. 2008). Analyses of historical changes in biomarker pigments from algae and phototrophic bacteria (cyanobacteria, purple sulfur bacteria) were used to: (1) quantify how phytoplankton abundance and community composition have varied in response to lake-level changes since ~ 1900; (2) determine if climate and lake-level change resulted in disproportionate increases in toxic cyanobacteria as is expected in warming, shallowing waters (Taranu et al. 2015; Bunting et al. 2016; Vogt et al. 2018) and (3) determine if any changes in lake physical structure (i.e., mixing regime, oxygenation, light penetration) due to historical lake-level variation are related to increased temporal variance of phytoplankton, such as seen in other prairie lakes which have undergone regime shifts (Carpenter et al. 2011; Bunting et al. 2016). Analysis of temporal patterns of primary producers may provide insights into the ecological effects of regional changes in lake levels (Dakos et al. 2015; Taranu et al. 2018) and will help forecast lake response to future

environmental change under a warmer and potentially more arid climate (Asong et al. 2016).

## Materials and methods

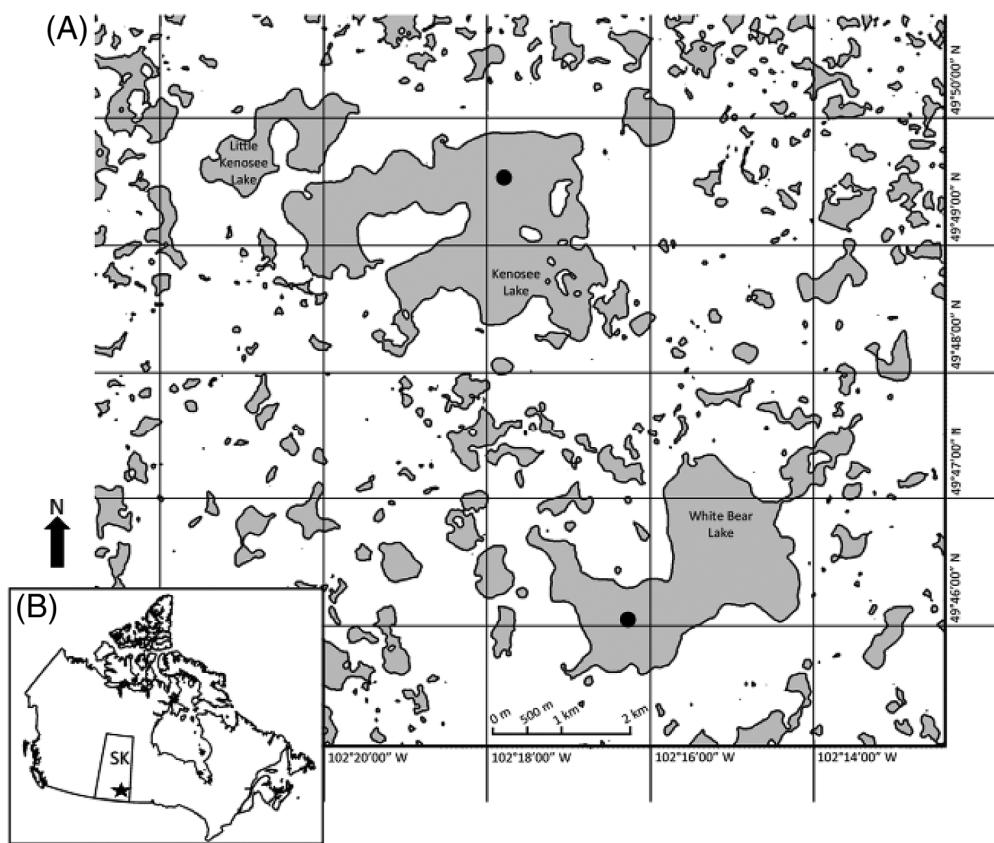
### Study sites

Kenosee Lake and White Bear Lake are located in the Moose Mountain Uplands of southeastern Saskatchewan, Canada (Fig. 1). These lakes are separated by ~ 2 km and share a humid, cold, continental climate (Köppen Dfb) with a mean annual temperature of 3.7°C and 427 mm yr<sup>-1</sup> of precipitation during the 1981–2010 period (Estevan, Saskatchewan, station data; Environment and Climate Change Canada [ECCC]; [https://climate.weather.gc.ca/climate\\_normals/results\\_1981\\_2010\\_e.html?searchType=stnProv&lstProvince=SK&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=2896&dispBack=0](https://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?searchType=stnProv&lstProvince=SK&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=2896&dispBack=0), accessed April 2021). Historical data suggest that while mean annual temperatures have increased by ~ 2°C over the past 80 yr, rainfall has increased only modestly and there has been little trend in snowfall (Supplemental Information Fig. S1; Homogenized Canadian Climate Station Data; ECCC; <https://www.canada.ca/en/environment-climate-change/service-s/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data.html>, accessed August 2021).

<https://www.canada.ca/en/environment-climate-change/service-s/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data.html>, accessed August 2021).

The two lakes are of similar size (~ 8–9 km<sup>2</sup>), although White Bear Lake is deeper ( $Z_{\max} = 15$  vs. 8 m) and drains a larger area (172 vs. 60 km<sup>2</sup>) than Kenosee Lake (Table 1; Vance et al. 1997; van der Kamp et al. 2008). The catchments of Kenosee and White Bear lakes exhibit poor hydrological integration and neither lake has channelized inflows or outflows. Instead, wetlands, sloughs, and lakes in the Moose Mountain Uplands rely on complex subsurface connections of saline, carbonate-rich groundwater that is eventually discharged on the adjacent prairie surface ~ 150 m below (Vance et al. 1997). Historically, Kenosee Lake spilled into White Bear Lake when water levels reached over 742 m above sea level (asl), but water conveyance between lakes has not been recorded since 1954 due to water-level low stands (see below) and the construction of a highway between the basins (Godwin et al. 2013).

Land use is similar within catchments of Kenosee and White Bear lakes, with ~ 55–57% cover by broadleaf deciduous forest primarily comprised of trembling aspen (*Populus tremuloides*), balsam poplar (*Populus balsamifera*), green ash



**Fig. 1.** (A) Map of Kenosee and White Bear Lakes with marked coring locations. Coring depths are 7.6 and 9.2 m for Kenosee Lake and White Bear Lake, respectively. Map interpreted from photograph courtesy of the U.S. Geological Survey (Landsat 8 OLI/TIRS database). (B) The location of Kenosee and White Bear lakes in the Moose Mountain Uplands of southeast Saskatchewan in relation to the Province of Saskatchewan (SK) in Canada.

**Table 1.** Summary of physical and chemical characteristics of Kenosee and White Bear lakes. Surface area ( $\text{km}^2$ ) and max depth (m) were obtained from van der Kamp et al. (2008), while physical (Secchi depth), chemical (total phosphorus [TP], total nitrogen [TN], total organic carbon [TOC]), salinity, pH, and Chl  $a$  were measured once per month from Jun to Sep in 2016 (mean  $\pm$  standard deviation).

Site	Kenosee Lake	White bear Lake
Surface area ( $\text{km}^2$ )	8	9
Drainage area ( $\text{km}^2$ )	60	172
Maximum depth (m)	8	15
Lake-level elevation (m asl)	741	729
TN ( $\mu\text{g}\cdot\text{L}^{-1}$ )	2048 $\pm$ 38	2533 $\pm$ 116
TP ( $\mu\text{g}\cdot\text{L}^{-1}$ )	27.50 $\pm$ 9.57	14.46 $\pm$ 10.41
TOC ( $\text{mg}\cdot\text{L}^{-1}$ )	26.75 $\pm$ 0.44	35.03 $\pm$ 1.97
Chl $a$ ( $\mu\text{g}\cdot\text{L}^{-1}$ )	9.44 $\pm$ 4.41	8.40 $\pm$ 7.79
Salinity ( $\text{g}\cdot\text{L}^{-1}$ )	1.07 $\pm$ 0.03	1.84 $\pm$ 0.06
pH	8.58 $\pm$ 0.14	8.72 $\pm$ 0.07
Secchi depth (m)	1.84 $\pm$ 1.36	2.58 $\pm$ 1.07

(*Fraxinus pennsylvanica*), white birch (*Betula papyrifera*), and Manitoba maple (*Acer negundo*; Henderson et al. 2002),  $\sim$  21–27% cover by other water bodies,  $\sim$  9–10% by grasses and shrubs, and < 1% agricultural cover (Agriculture and Agri-Food Canada 2013). The catchments of Kenosee and White Bear lakes have never been cleared. A portion of the catchment of White Bear Lake is encompassed by White Bear First Nations reserve created in 1875, while the remaining catchment area of both lakes became part of a Canadian federal forest reserve in 1894 and a Saskatchewan provincial park in 1931 (Henderson et al. 2002). Despite these designations and protections, significant recreational development has occurred on the shorelines of both lakes since the 1960s including the construction of cottages and golf courses. Monthly sampling during June–September 2016 showed that both lakes are currently hyposaline, alkaline, and mesotrophic (Table 1). Currently, both basins do not stratify, although nothing is known of the interannual variation in mixing intensity.

### Lake-level history

Annual lake-level data for Kenosee and White Bear lakes were recorded by the Government of Canada (Historical Hydrometric Data; [https://wateroffice.ec.gc.ca/mainmenu/historical\\_data\\_index\\_e.html](https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html), accessed April 2021) between 1964 and 2016. Values for White Bear Lake between 1910 and 1964 were obtained from Cullimore and Griffin (1979). While no data are available regarding Kenosee Lake's water-levels prior to 1964, a strong linear relationship between lakes levels since 1964 ( $R^2 = 0.72$ ,  $p < 0.0001$ ) suggest that both basins would have experienced similar degrees of lake-level variation during the early 20<sup>th</sup> century.

### Paleolimnological analyses

Sediment cores were collected from deep-water sites from Kenosee and White Bear lakes using a Glew gravity corer (Glew 1989) in August 2016 (Fig. 1). The Kenosee Lake core was collected at  $\sim$  7.6 m depth ( $49^\circ 49.455'\text{N}$ ,  $102^\circ 18.882'\text{W}$ ) and was  $\sim$  57 cm in length, while the White Bear Lake core was taken at  $\sim$  9.2 m depth ( $49^\circ 49.455'\text{N}$ ,  $102^\circ 18.882'\text{W}$ ) and was  $\sim$  56-cm long. Both cores were sectioned on site at 0.5-cm intervals, stored in the dark on ice during transport, and refrigerated until analysis within 4 months of collection. Sediments from the top 40 cm of each core were freeze-dried (72 h, 0.1 Pa) for subsequent analyses of  $^{210}\text{Pb}$  and  $^{137}\text{Cs}$  activities, stable isotope content (carbon [C], nitrogen [N]), and pigment biomarker concentrations in the Institute of Environmental Change and Society at the University of Regina.

Sediment chronology was based on  $^{210}\text{Pb}$  and  $^{137}\text{Cs}$  activities quantified using gamma spectrometric analysis of 12 evenly spaced sections of each core (Appelby et al. 1986). Sediment age and mass accumulation rates ( $\text{g}\text{ cm}^{-2}\text{ yr}^{-1}$ ) were calculated using the constant rate of supply (CRS) model (Binford 1990). Sediment age-depth relationships were refined using shape-constrained additive models (SCAMs) with monotone decreasing P-splines via the *scam* package (Pya 2021) with generalized cross-validation smoothness parameter selection in R (R Core Team 2021).

Whole dried sediments were analyzed for stable isotope ( $\delta^{15}\text{N}$ ,  $\delta^{13}\text{C}$ ) and elemental content (N%, C%) by combustion using a Thermoquest Delta Plus isotope ratio mass spectrometer equipped with a Thermoquest NC2500 elemental analyzer (Savage et al. 2004). Carbon and nitrogen isotope values were standardized against international standards (Pee Dee Belemnite and atmospheric  $\text{N}_2$ , respectively) and expressed using standard ‰ notation. Elemental composition of whole sediments were estimated as % dry mass for N (N%) and C (C%) content, and were used to estimate C : N mass ratios.

High-performance liquid chromatography (HPLC) was used to quantify fossil pigment concentrations from alternate sediment sections in the Kenosee and White Bear cores following Leavitt and Hodgson (2001). Pigments were extracted from 15 to 100 mg of freeze-dried sediments by an 80 : 15 : 5 (by volume) solution of HPLC-grade acetone, methanol, and water. Extracts were filtered (0.22-μm pore) and evaporated under inert  $\text{N}_2$  gas, before being redissolved into injection solution. Concentrations of fossil pigments were measured using an Agilent model 1260 HPLC calibrated with authentic pigment standards and using Sudan II as an internal reference. Pigment interpretation followed Leavitt and Hodgson (2001) with concentrations of chlorophyll  $a$  (Chl  $a$ ) derivative pheophytin  $a$ , and  $\beta$ -carotene used as indicators of total phototroph abundance (Leavitt and Hodgson 2001). Other taxon-specific pigments included fucoxanthin (siliceous algae), diatoxanthin (primarily diatoms), alloxanthin

(cryptophytes), pheophytin *b* (chlorophytes), echinenone (total cyanobacteria), and canthaxanthin (Nostocales cyanobacteria). Lutein and zeaxanthin could not be separated and were combined as indicators of bloom-forming taxa (Leavitt and Hodgson 2001). In addition, okenone was used as an indicator of purple sulfur bacteria (Leavitt et al. 1989; Leavitt and Hodgson 2001). All pigment concentrations were expressed as nmoles pigment g<sup>-1</sup> carbon (Leavitt et al. 1994). The ratio of Chl *a* to pheophytin *a* (Chl : pheo) was used as a metric of changes in preservation environment (Leavitt and Hodgson 2001), while the ratio of UVR-absorbing scytonemin derivatives to the sum of other carotenoids (alloxanthin, lutein-zeaxanthin, diatoxanthin) was used as an index of past exposure to UV irradiance (Leavitt et al. 1997).

### Numerical analyses

Temporal trends in pigment and other geochemical proxies were estimated using generalized additive models (GAMs) using the *mgcv* package (Wood 2011, 2017; Simpson 2018). Specifically, pigment concentrations were estimated using a location-scale hierarchical GAM (HGAM) where both the mean and scale predictors used a global smooth of year and a factor smooth for each combination of the 2 lakes and

10 pigments for a total of 20 factors (model GS in Pedersen et al. 2019). The global smooth accounted for the common trend between both lakes and all pigments, while the factor smooth accounted for the deviations at the pigment and lake level from the global smooth. The model deviations were fit assuming a common smoothness parameter between lakes and pigments, but do not account for common trends between lakes or pigments separately. Both the global smooth and the factor smooth were fit using cubic regression splines. Finally, the scale predictor also accounted for the period of time represented by each core slice to account for changes in temporal averaging between adjacent samples. This was facilitated with the addition of a smooth of each sample's log-transformed temporal interval and by fitting the smooth with adaptive splines.

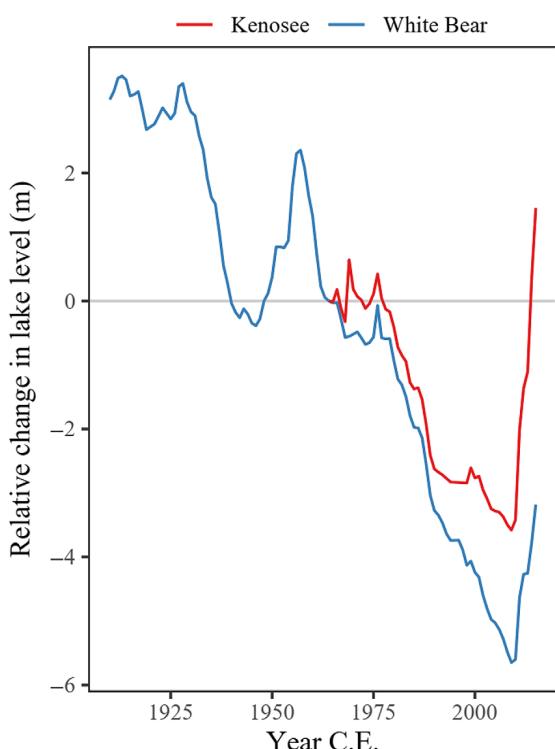
Pigment variances were extracted from pigment concentrations by calculating the product of the mean and shape estimates from the concentration HGAM. Credible intervals (95%) for the variance values were obtained by running 10,000 simulations and taking the 2.5% and 97.5% quantiles of the posterior distributions. Resultant pigment variances were also modeled using an HGAM under the same parameters as described above. In addition, Chl : pheo ratios and UV indexes were modeled individually using HGAMs with Gamma and Tweedie distributions, respectively. Both models used a smooth for year and lake, such that each lake had a different smoothness parameter (model *I* in Pedersen et al. 2019). In all models, observations were weighted by temporal resolution and the smoothness parameter was estimated using a restricted maximum likelihood approach (Simpson 2018). To identify periods of significant change, the first derivative of the estimated smooth trend was evaluated from the relevant model of each proxy (Bunting et al. 2016; Simpson 2018). Here, the first derivative of each proxy smooth was estimated using the *gratia* package in R (Simpson 2021). Periods of significant change were identified where the 95% credible interval on the estimated derivative excluded 0.

All statistical analyses were performed in the R statistical environment (R Core Team 2021). The *tidy* and *dplyr* packages (Wickham 2021; Wickham et al. 2021) were used for data wrangling, while plots were created using the *ggplot2* and *cowplot* packages (Wickham et al. 2016; Wilke 2020). Code for analyses is available on GitHub at <https://github.com/simpson-lab/kenosee-white-bear>.

## Results

### Lake-level

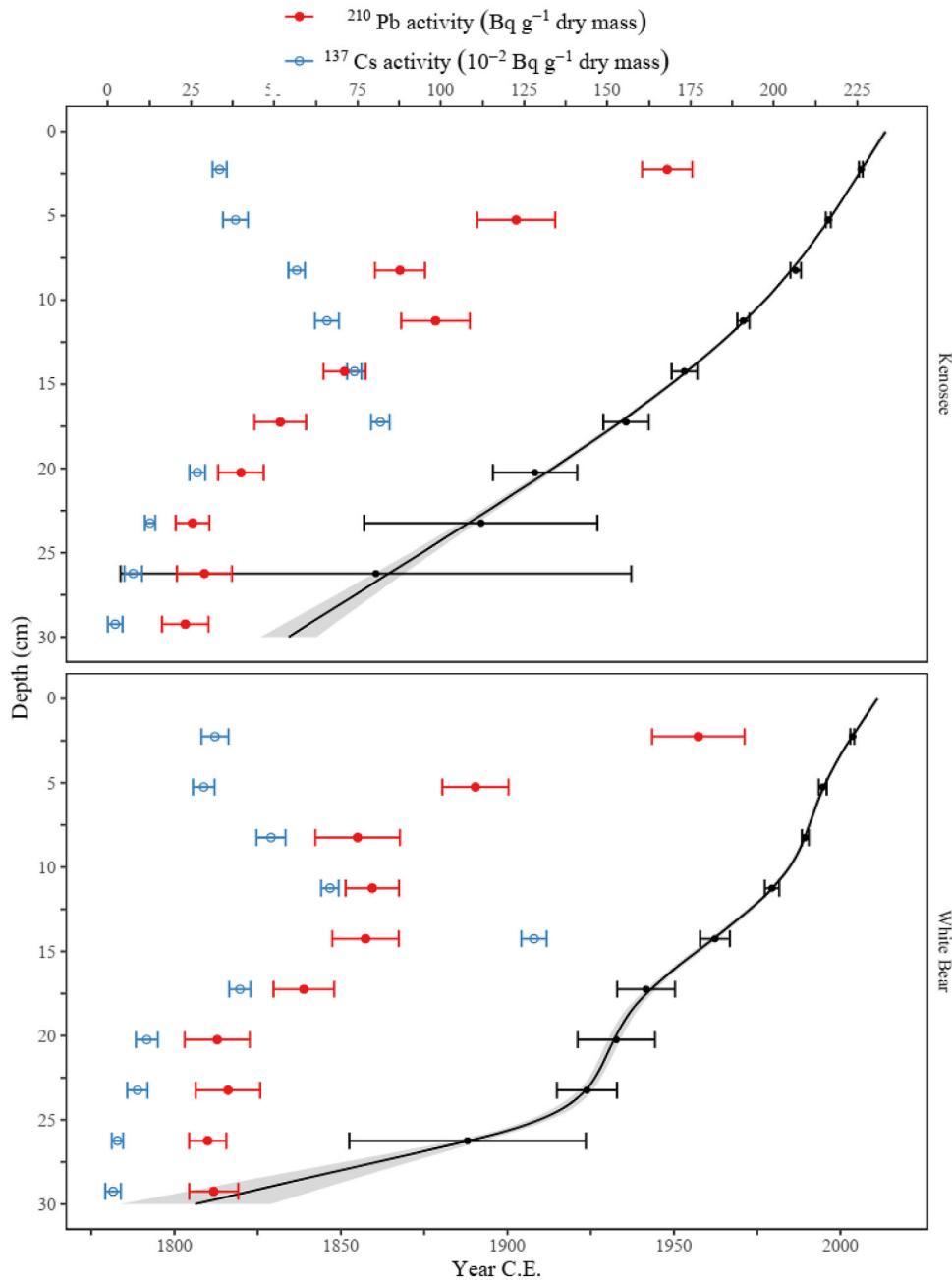
Historical records suggested that both Kenosee and White Bear lakes experienced water-level variation of > 8 m since the early 20<sup>th</sup> century (Fig. 2). Taking the first common year of record (1964) as a benchmark, water levels in White Bear Lake were ~ 4 m higher between ~ 1910 and ~ 1930 before



**Fig. 2.** Water levels of Kenosee Lake (red) and White Bear Lake (blue) relative to their levels in 1964. Data from 1964 to 2015 were obtained from the Government of Canada Water Office; data prior to 1964 for White Bear Lake were obtained from estimates from Cullimore and Griffin (1979).

declining toward a plateau by 1950. While lake levels rose  $\sim 2$  m during the 1950s, values declined again to the late 1960s. The water levels of both lakes were stable until  $\sim 1975$  when marked declines of up to 6 m and 3 m occurred in White Bear and Kenosee lakes, respectively. Water levels reached a minimum at  $\sim 2010$  ( $-4$  to  $-5.5$  m) before rapidly rising by 2 m in White Bear Lake and 4 m in Kenosee Lake in recent years (Fig. 2). Given the strong

correlation between Kenosee and White Bear lake levels since 1964 ( $R^2 = 0.72$ ,  $p \leq 0.0001$ ), it is likely that Kenosee Lake experienced similar water-level fluxes prior to 1964. Together, these findings suggest that Kenosee and White Bear lakes have experienced water-level variations equivalent to  $\sim 75\%$  and  $\sim 55\%$  of their present depth, respectively, with only very recent increases toward historical benchmarks (Fig. 2).



**Fig. 3.** Activities of  $^{210}\text{Pb}$  and  $^{137}\text{Cs}$  with associated error estimates ( $1\sigma$ ) by core depth for Kenosee Lake and White Bear Lake. Age-depth relationships were estimated using SCAMs (shape-constrained additive models)-based CRS (constant rate of supply) models of  $^{210}\text{Pb}$  activity in each core. Inferred dates are also presented with error ranges ( $1\sigma$ ) by core depth for Kenosee Lake and White Bear Lake.

### Sediment chronology

Activity of  $^{210}\text{Pb}$  declined with sediment depth in the Kenosee and White Bear cores with little evidence of sediment mixing (Fig. 3). Activity profiles of  $^{137}\text{Cs}$  were well defined in White Bear Lake sediments, with a clear maximum in  $^{210}\text{Pb}$ -dated intervals corresponding to peak atmospheric nuclear testing in 1963 at 14 cm (Fig. 3). The  $^{137}\text{Cs}$  peak was less well defined in Kenosee Lake, with a maximum at  $\sim 17$  cm. Age-depth models suggested that bulk dry sediment accumulation rates were comparable between sites, whereas SCAMs based on the CRS models suggest that ages at  $\sim 30$  cm were essentially the same;  $\sim 1830$  and  $\sim 1810$  for Kenosee and White Bear lakes, respectively (Fig. 3).

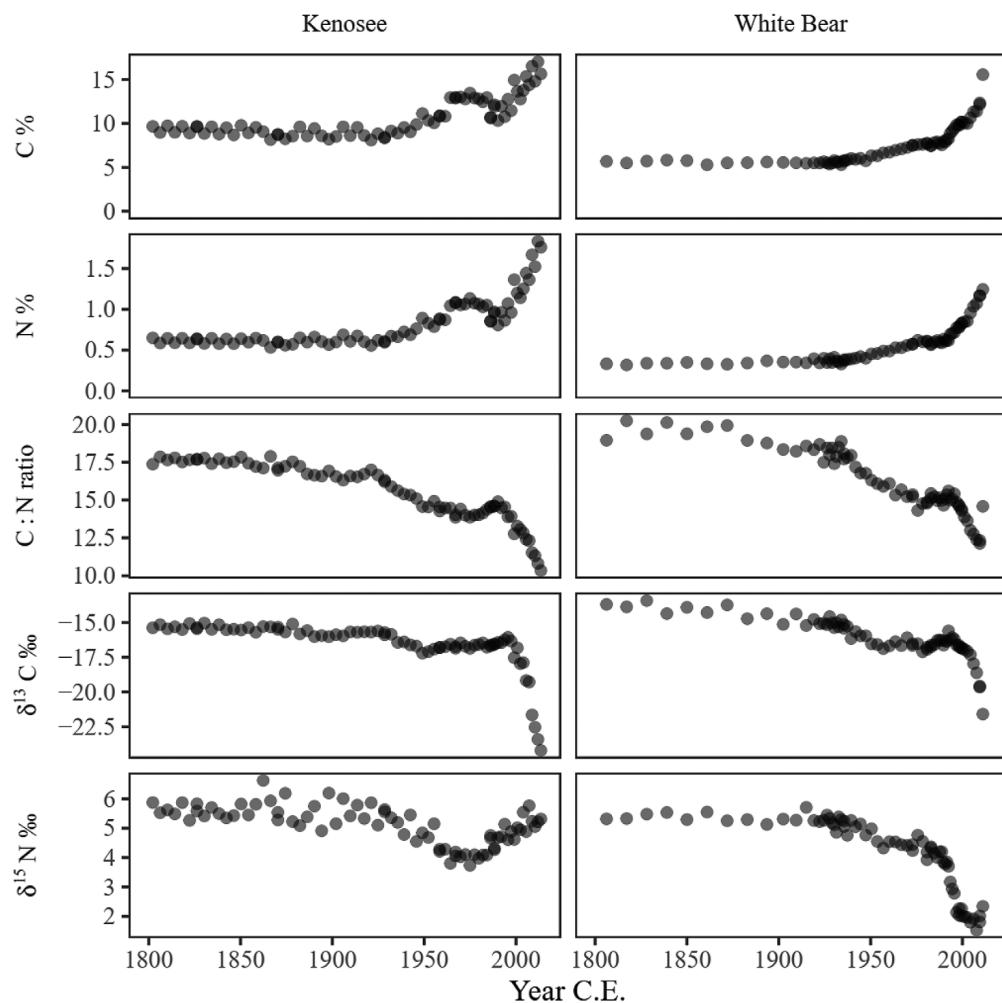
### Geochemistry and stable isotopes

Geochemical trends in stable isotope values were generally similar in the cores from Kenosee and White Bear lakes (Fig. 4). In both cores, C and N content (% by mass) was low prior to  $\sim 1930$ , but rose rapidly afterward to a transient

plateau ca. 1950–1975, before continuing to historical maxima in the most recently deposited sediments. The C : N ratios of both lakes exhibited an inverse relationship to C and N content, with stable values of  $\sim 18$  prior to 1900, declining to a plateau before accelerating to a minimum after ca. 2000 (Fig. 4). Sedimentary  $\delta^{13}\text{C}$  values were relatively stable and elevated at both sites before  $\sim 2000$  ( $\sim 15\text{\textperthousand}$ ), after which isotope values declined to  $\sim -22.5\text{\textperthousand}$  in both lakes (Fig. 4). In contrast, trends in  $\delta^{15}\text{N}$  were markedly different between lakes. In Kenosee Lake,  $\delta^{15}\text{N}$  ratios decreased from  $\sim 6\text{\textperthousand}$  to  $\sim 4\text{\textperthousand}$  after approximately 1925 before returning to more enriched values after  $\sim 1960$  (Fig. 4), whereas in White Bear Lake,  $\delta^{15}\text{N}$  ratios continue to decline after  $\sim 1960$  with depletions to  $\sim 2\text{\textperthousand}$  in sediments deposited since ca. 2000 (Fig. 4).

### Phototrophic pigments

Analysis of sedimentary carotenoid and chlorophyll pigments revealed significant changes in the composition, concentration, and variance of primary producers over the

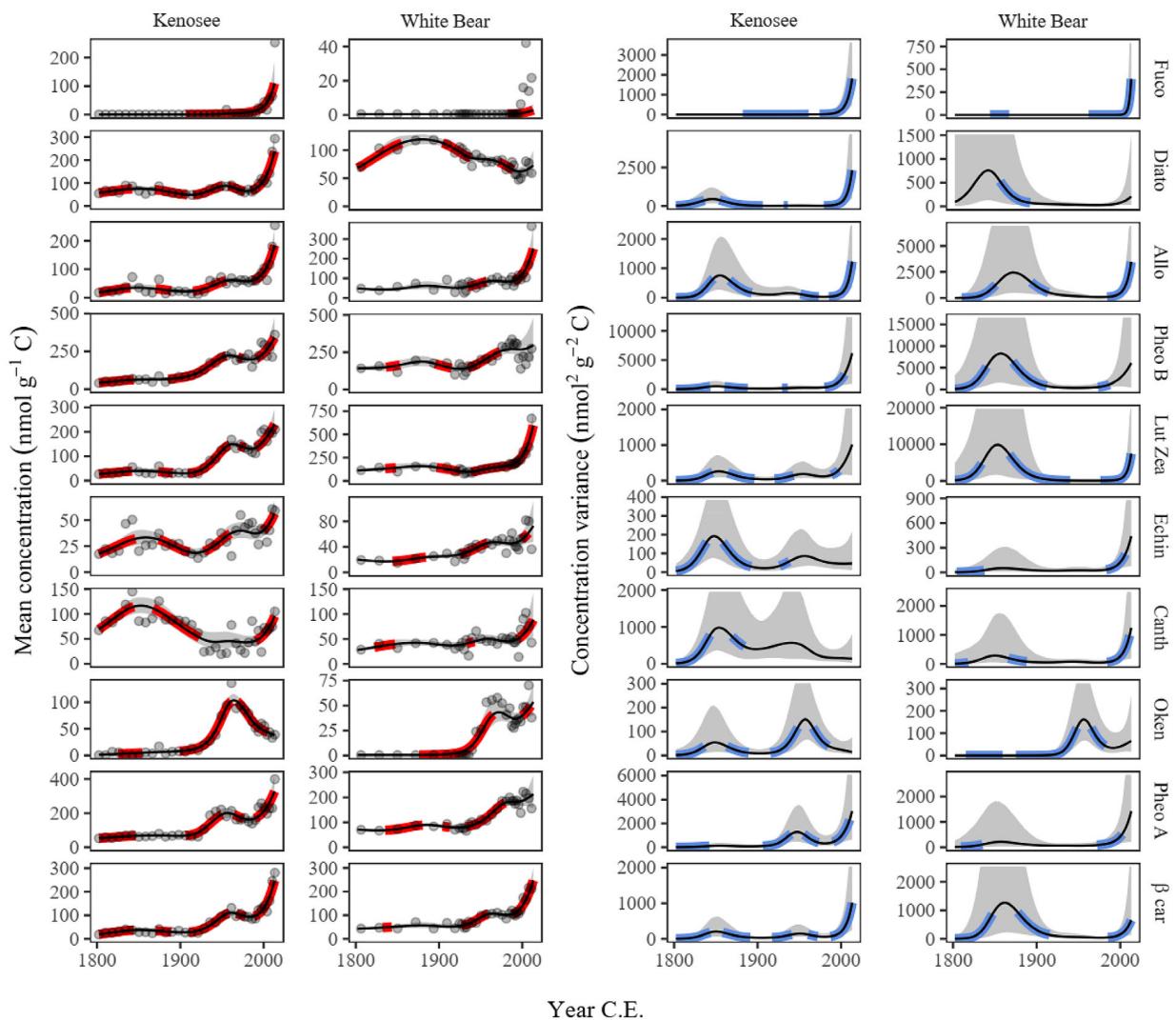


**Fig. 4.** Carbon (C%) and nitrogen (N%) content, C : N ratios, and stable isotope ratios of  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  (‰) plotted by year for Kenosee Lake and White Bear Lake.

past ~200 yr of Kenosee and White Bear lakes (Fig. 5). In these analyses, significant changes refer to periods of time when the slope (i.e., the first derivative) of pigment concentrations and variances are statistically significant from 0 (see the Methods section). In Kenosee Lake, concentrations of pigments indicative of siliceous algae (fucoxanthin), diatoms (diatoxanthin), cryptophytes (alloxanthin), chlorophytes (pheophytin *b*), and other bloom-forming taxa (lutein-zeaxanthin) were generally low, but did show periods of significant increase prior to ~1900. Abundances of chlorophytes (pheophytin *b*) increased significantly after ~1900, with the largest changes occurring after ~1930, concomitant with lake-level fall, whereas densities of diatoms (diatoxanthin),

cryptophytes (alloxanthin), and bloom-forming taxa (lutein-zeaxanthin) increased significantly across the 20<sup>th</sup> century with periods of minor declines centered at ~1975 and additional decreases in lake level (Figs. 2, 5). Trends in these phytoplankton were similar in White Bear Lake sediments, with the exception of a slightly delayed onset of significant increases at ~1930 (Fig. 5). In that lake, diatoxanthin exhibited distinctly high concentrations prior to 1900, but declined throughout the 20<sup>th</sup> century with statistically significant intervals of decline at ~1930 and ~1970 (Fig. 5).

In Kenosee Lake, concentrations of pigments from cyanobacteria (echinenone and canthaxanthin) were relatively elevated at ~1800 and rose significantly to ~1850 before

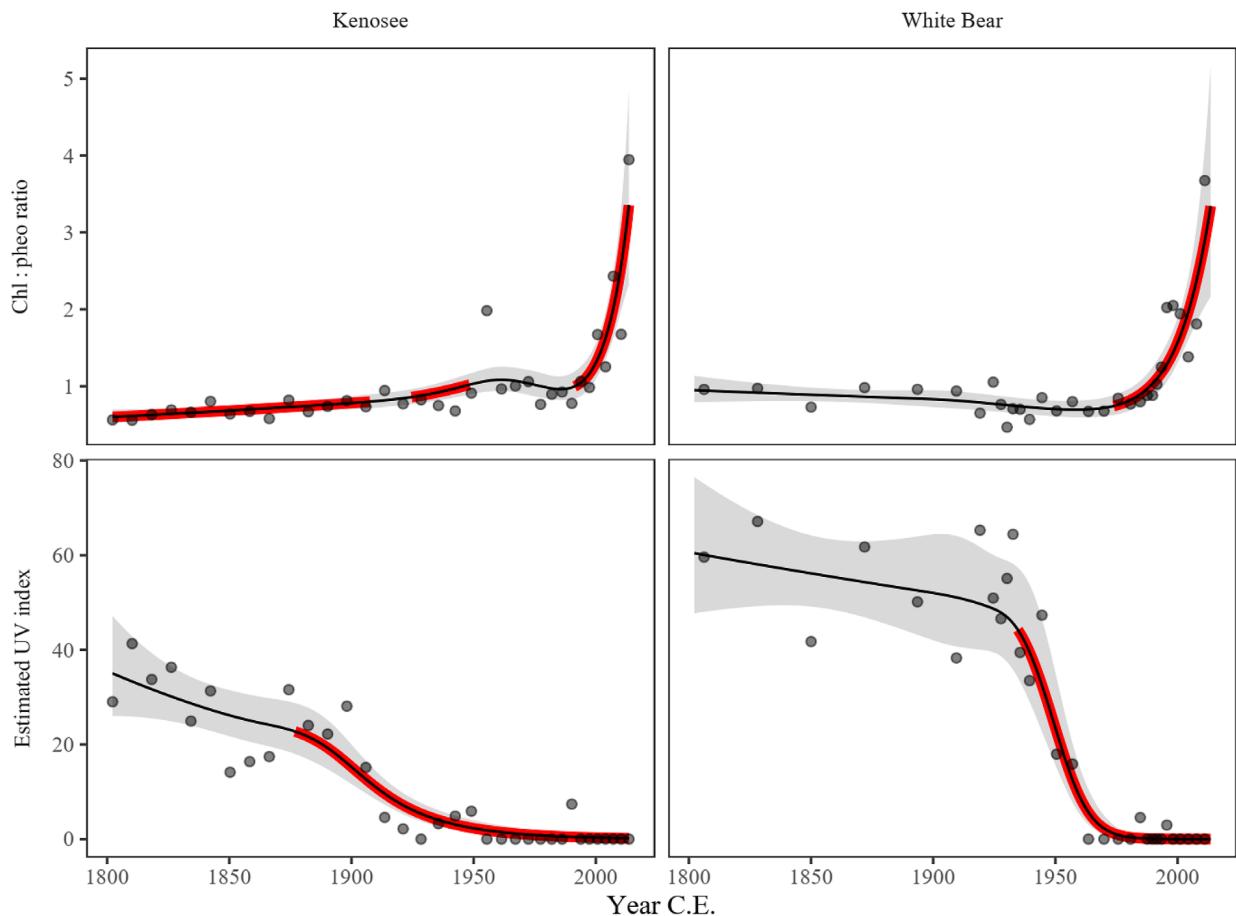


**Fig. 5.** Mean concentrations (left) and concentration variances (right) of fossil pigments by year in sediments from Kenosee Lake and White Bear Lake. Pigment mean concentrations and concentration variances are fitted with hierarchical generalized additive models (HGAMs). In all plots, solid lines are the fitted model trends, gray shading represents 95% confidence intervals of the trends, and emboldened sections of the trends represent significant changes in pigment mean concentration or concentration variance. Fuco = fucoxanthin (siliceous algae), Diato = diatoxanthin (primarily diatoms), Allo = alloxanthin (cryptophytes), Pho B = pheophytin *b* (chlorophytes), Lut Zea = lutein-zeaxanthin (chlorophytes and cyanobacteria pigments, i.e., “bloom-forming” taxa), Echin = echinenone (total cyanobacteria), Canth = canthaxanthin (*Nostocales* cyanobacteria), Oken = okenone (purple sulfur bacteria), Pho A = pheophytin *a* (total production), β car = β-carotene (total production).

declining significantly to 1900 (Fig. 5). Abundance of total cyanobacteria (echinenone) then rose significantly from ~ 1930 to 1950 alongside falls in lake level (Figs. 2, 5). In contrast, potentially N<sub>2</sub>-fixing colonial cyanobacteria (canthaxanthin) declined throughout the early 20<sup>th</sup> century, reaching stable and relatively low concentrations between ~ 1930 and 2000 before rising significantly after ~ 2000 (Fig. 5). Concentrations of both pigments were low in White Bear Lake prior to 1900, despite periods of significant increases during this time, and slowly rose over the 20<sup>th</sup> century, with significant increases in echinenone at ~ 1930–1950 alongside declines in lake level (Figs. 2, 5). In both lakes, concentrations of okenone from obligately anaerobic purple sulfur bacteria were negligible prior to ~ 1900, but increased substantially during the early 20<sup>th</sup> century and reached maxima at ~ 1950, a period of stable water levels at both sites (Figs. 2, 5). Thereafter, concentrations of okenone declined significantly in Kenosee lake during periods of lake-level increase (~ 1950–1970) and subsequent decline (~ 1970–2010) (Figs. 2, 5). These declines in okenone concentration did not reach pre-1900 minimum values. In White Bear Lake, concentrations of okenone

plateaued during ~ 1950–2000 despite substantial lake-level changes, before rising significantly thereafter (Figs. 2, 5). In addition, overall biomarkers of total phototroph production (pheophytin *a*, β-carotene) were relatively stable before ~ 1900, but increased significantly between ~ 1900 and 1950 in Kenosee Lake and between ~ 1930 and 1970 in White Bear Lake (Fig. 5). Maximum values of these pigments occurred in both lakes after ~ 2000, coeval with recent increases in lake levels (Figs. 2, 5).

Temporal trends in variance of primary producers were similar in Kenosee and White Bear lakes sediments (Fig. 5). At both sites, variance of diatoms (diatoxanthin), cryptophytes (alloxanthin), and green algae (pheophytin *b*, lutein) was elevated during the mid-19<sup>th</sup> century, centered at ~ 1840–1860, before declining significantly to low values by 1900 (Fig. 5). Variance measures of these phytoplankton remained low through the period of lake-level variation during the 20<sup>th</sup> century before increasing significantly after ~ 2000 (Figs. 2, 5). In Kenosee Lake, variance in cyanobacterial abundance (echinenone, canthaxanthin) increased significantly prior to ~ 1850 then significantly declined to low values by 1900



**Fig. 6.** Precursory Chl *a* to derivative pheophytin *a* (Chl : pheo) ratios and estimated UV index values plotted by year for Kenosee Lake and White Bear Lake. In all plots, solid lines are the fitted model trends, gray shading represents 95% trend confidence intervals, and emboldened sections represent periods of significant change.

(Fig. 5). In White Bear Lake, 19<sup>th</sup> century changes in variance of these pigments were much less marked, although still significant increases in cyanobacterial variance were recorded up to ~1850 (Fig. 5). Similarly, variance in purple sulfur bacteria (okenone) differed between lakes before ~1900, with slightly elevated values in Kenosee in the mid-19<sup>th</sup> century, but consistently low variance in White Bear Lake during the same interval (Fig. 5). Changes in okenone variance were more coherent between lakes after ~1900, rising significantly between ~1930 and ~1950 during the period of lake-level decline (Figs. 2, 5). These peaks in okenone variance were short-lived and declined significantly to ~1975, a period of substantial lake-level variation. In contrast to other pigments, variation in okenone time series did not rise again during late 20<sup>th</sup> century periods of lake-level decline, nor did it increase significantly when lakes refilled after ~2010 (Figs. 2, 5). Finally, there were few common patterns of historical change in variation of total primary producers (pheophytin *a*, β-carotene), with a higher amount of significant variance changes across the 20<sup>th</sup> century in Kenosee Lake and more significant changes during the 19<sup>th</sup> century at White Bear Lake (Fig. 5). In both basins, variance of these pigments rose significantly after ~2000 (Figs. 2, 5).

Analysis of changes in preservation environment (as Chl : pheo ratios) suggested that there was little variation in sedimentary pigment preservation in either lake until the most recently deposited sediments (Fig. 6). Although changes in Chl : pheo ratios were significant in Kenosee Lake for the first 150 yr of the record, these changes were minor relative to those seen after ~2000. Reconstruction of the UVR index suggested that phytoplankton were exposed to relatively high levels of UV radiation during the 19<sup>th</sup> century, but that exposure declined significantly through the 20<sup>th</sup> century, culminating in minimum values in the recent sediments of both lakes (Fig. 6).

## Discussion

Kenosee and White Bear lakes have experienced lake-level variability of >8 m during the last century (Fig. 2), likely reflecting rising regional temperatures, increased evaporation, and complex interactions between Arctic, Gulf, and Pacific air masses which affect regional snowpack, spring runoff, and groundwater availability (Bonsal et al. 2006; Pomeroy et al. 2007; Liu et al. 2008; van der Kamp et al. 2008; McCullough et al. 2012). Lake-level decline was marked by coeval increases in primary production beginning ca. 1930, including elevated abundance of total phytoplankton, diatoms, cryptophytes, and chlorophytes (Fig. 5), concomitant increases in sedimentary C and N content (Fig. 4) and bulk sedimentation rates (Fig. 3). However, despite increasing temperatures and loss of over 50% of lake depth, neither total nor colonial diazotrophic cyanobacteria increased substantially during the 20<sup>th</sup> century relative to prior levels (Fig. 5). Instead,

fossil concentrations of okenone from anaerobic purple sulfur bacteria increased in both lakes during ~1930–1950 (Fig. 5), showing that light penetrated to anoxic habitats during this interval (Pfennig 1978; Leavitt et al. 1989). While variance in total primary production did not increase substantially prior to the development of illuminated anoxic habitats during intervals of stable and low lake level, temporally coherent and significant increases in okenone variance occurred during lake-level declines between ~1930 and 1950 suggesting that the development of illuminated and anoxic habitats in both Kenosee and White Bear lakes may represent a regime shift with respect to bacteria and physical lake structure (Scheffer et al. 2001; Dakos et al. 2015). These patterns are consistent with the establishment of bacterial plates within transient chemoclines following lake-level decline and evaporative concentrations of solutes in deep waters (Leavitt et al. 1989; Vinebrooke et al. 1998; Pham et al. 2009). This novel physical lake structure does not appear to represent an alternate stable state, as subsequent and substantial changes in lake-level restricted the anoxic and illuminated environments and suggests that any possible regime shift was transient and lacked hysteresis (Dakos et al. 2015; Ratajczak et al. 2018). Regardless, these abrupt changes in lake level, mixing, oxygenation, light penetration, and biotic structure may become more common in the Northern Great Plains, where future climate change is expected to increase evaporative forcing by 2050 due to a ~5°C warming and only a modest (<5%) increase in precipitation (Asong et al. 2016; Sauchyn et al. 2020).

## Climate forcing of upland lakes

Lake levels at Kenosee Lake and White Bear Lake declined >8 m since 1910 (Fig. 2). While extraordinary relative to many boreal lakes, such large excursions are common in the Northern Great Plains (Fritz 1990; van der Kamp et al. 2008) due to water deficits (precipitation-evaporation) ranging -20 to -60 cm yr<sup>-1</sup> (Pham et al. 2009; Haig et al. 2020, 2021), and pronounced shifts in the relative importance of winter air masses that control the hydrodynamics of snow accumulation and melting (Bonsal et al. 2006; Pomeroy et al. 2007; Liu et al. 2008). In the Canadian Prairie region, >70% of annual precipitation falls as summer rain, but up to 80% of regional runoff is derived from spring snow melt (Akinremi et al. 1999). This snowmelt recharges surface and intermediary-depth aquifers (van der Kamp and Maathuis 1991) that can also contribute water to some prairie and parkland lakes (Shaw and Prepas 1990). In general, regional precipitation is regulated by a complex interplay between air masses from the Arctic, Pacific Ocean, and Gulf of Mexico (Bryson and Hare 1974) that is additionally influenced by the El Niño-Southern Oscillation, North Atlantic Oscillation, and Pacific Decadal Oscillation climate systems (Trenberth and Hurrell 1994; Hurrell 1995; Mantua et al. 1997). Paleoclimate and modeling analyses reveal that centennial- and continental-scale variation in these air masses

affect the position of the winter jetstream and, in turn, the precipitation supply which creates decadal-scale changes in runoff, lake level, and salinity (Fritz 1990; Michels et al. 2007; Liu et al. 2008; McCullough et al. 2012). Relationships between runoff and lake level are further complicated by the presence of numerous shallow water bodies within lake catchments (Fig. 1) that variously “fill and spill” depending on antecedent meteorological conditions (Coles and McDonnell 2018; Haig et al. 2021). Finally, land-use practices can alter lake levels by channelizing surface flow, removing wetlands, extracting for human use, modifying forests, or through agricultural irrigation (Fang et al. 2007; Mao and Cherkauer 2008).

Patterns of hydrology in White Bear and Kenosee lakes (Fig. 2) are consistent with known variability in regional climate and land-management practices (Vance et al. 1997; Vinebrooke et al. 1998). For example, stable elevated lake levels were recorded in White Bear Lake until onset of the prolonged prairie droughts of the 1920s and 1930s when blocking high pressure cells reduced the influx of moisture from the Gulf of Mexico (Bonsal et al. 2006). Lake levels recovered briefly during the 1950s before declining again to a plateau until the early 1970s. These events correspond to rapid changes between arid and pluvial conditions during the 1950s and 1960s (Henderson et al. 2002). After ~1970, lake levels declined continuously until the early 21<sup>st</sup> century, reflecting a both 0.95 cm yr<sup>-1</sup> decline in prairie winter precipitation during this interval (Akinremi et al. 1999; Henderson et al. 2002) and local water management practices which included water extraction for nearby golf courses and the construction of a highway that restricted flow between adjacent waterways, including Kenosee and White Bear lakes (Godwin et al. 2013). Finally, sudden recent increases in lake level of 2–4 m may reflect changes in the supply of moisture into the region, similar to the step-change increase in precipitation and runoff in nearby Manitoba during the 1990s (McCullough et al. 2012; Dumanski et al. 2015). In this case, delays in lake-level response (Fig. 2) may arise because antecedent arid conditions created substantial hydrological sinks in both lakes’ catchments by lowering water levels in many small water bodies (Fig. 1). The presence of empty hydrological sinks decoupled precipitation and runoff until the sinks were filled and regained a steady state (“fill and spill”) exchange of influx and outflow of water (Coles and McDonnell 2018; Haig et al. 2021), although further local research is needed to validate this mechanism.

#### Phototroph response to lake-level declines

Analysis of chemically stable algal and cyanobacterial pigments revealed similar increases in primary production in Kenosee and White Bear lakes during the 20<sup>th</sup> century (Fig. 5). Specifically, when analyzed using GAMs, mean concentrations of biomarkers for diatoms (diatoxanthin), cryptophytes (alloxanthin), chlorophytes (pheophytin *b*, lutein), total

cyanobacteria (echinenone), and all primary producers (pheophytin *a*, β-carotene) increased significantly from ~1930 until the 21<sup>st</sup> century, although onset of eutrophication was delayed in White Bear Lake, compared to signals from Kenosee Lake (Fig. 5). Elevated lake production is consistent with trends seen in lowland lakes throughout the Canadian Prairies and normally reflect substantial changes in land use and nutrient influx (Leavitt et al. 2006; Pham et al. 2008; Maheaux et al. 2016). In contrast, catchments of upland Kenosee and White Bear lakes have not been subject to extensive modification, beyond shoreline development, suggesting that more extensive algal growth arose because of marked lake-level declines after ~1930 and ~1975 (Fig. 2). Although speculative, we note that shallow lakes are usually more productive than deeper basins of given size (Jeppesen et al. 2014) due to a higher fraction of profundal sediments in contact with warm surface waters and consequently elevated rates of internal nutrient loading (Søndergaard et al. 2013). Furthermore, phytoplankton production in prairie lakes can increase when evaporation concentrates epilimnetic nutrients and other solutes (Oduor and Schagerl 2007; Wissel et al. 2011; Vogt et al. 2018).

Historical variation in geochemical and isotopic sediment features were generally comparable in Kenosee and White Bear lakes and were consistent with pigment inferences of increased lake production after ~1930 (Figs. 3, 4). Specifically, variation in C and N content were inversely correlated with White Bear Lake water levels ( $R^2_{adj} = 0.76, 0.89$ ), whereas C : N mass ratios varied with lake level until the ~2010 minimum. Similarly, bulk sedimentation in White Bear Lake, and to a lesser extent Kenosee Lake, exhibited marked acceleration during intervals of declining lake level in the 20<sup>th</sup> century, with more modest rates of accumulation between ~1940 and 1970 when lake levels were generally stable beyond a transitory 2 m change during the 1950s (Fig. 3). Together, these patterns are consistent with increased deposition of organic matter resulting from elevated autochthonous primary production (Meyers and Teranes 2001), as recorded by coeval declines in C : N ratios from values characteristic of terrestrial plants (C : N ~ 20–25) to those associated with autochthonous organic matter (C : N ~ 8–12; Gu et al. 2006; Woodward et al. 2012). Depleted  $\delta^{13}\text{C}$  values in both lakes are also consistent with elevated in situ primary production during the 20<sup>th</sup> century, reflecting increased photosynthetic uptake of isotopically depleted CO<sub>2</sub> from respiration in situ sources (Meyers and Teranes 2001; Woodward et al. 2012) as seen in other eutrophied prairie lakes (Bunting et al. 2016).

Unlike most biogeochemical proxies, late 20<sup>th</sup> century measures of nitrogen cycling (as  $\delta^{15}\text{N}$ ) differed between Kenosee and White Bear lakes, with sedimentary enrichment at the former site and depletion at the latter (Fig. 4). In general, historical patterns of  $\delta^{15}\text{N}$  were correlated strongly with concentrations of okenone from purple sulfur bacteria ( $R^2_{adj} = 0.79, p < 0.0001$ ), taxa known to fix nitrogen in

illuminated anaerobic habitats (Madigan 1995). We infer that cyanobacteria did not contribute substantially to fixed N pools (c.f., Hayes et al. 2019), as concentrations of canthaxanthin from potentially diazotrophic cyanobacteria (Leavitt and Hodgson 2001; Hayes et al. 2019) were uncorrelated to historical variation in  $\delta^{15}\text{N}$  values (Figs. 3, 5). Thus, while it is possible that enhanced shoreline development since the 1960s (e.g., cottages and golf courses) may have added isotopically enriched N from fertilizers or wastes (Botrel et al. 2014), the absence of common patterns in  $\delta^{15}\text{N}$  of the lakes since the 1960s is more consistent with differential supply of fixed N from diazotrophic purple sulfur bacteria.

Historical variation in *Nostocales* cyanobacteria (canthaxanthin) did not show a close correspondence to either observed changes in lake levels (Fig. 2) or other sediment proxies of lake production (Figs. 3–5), despite cyanobacterial preference for warm, nutrient-rich, shallow conditions (Paerl and Paul 2012; Vogt et al. 2018), their presence in the current phytoplankton (Bos et al. 2019), and fossil concentrations which were similar those in other regional eutrophic lakes (Leavitt et al. 2006; Bunting et al. 2016; Maheaux et al. 2016). For example, *Nostocales* were most abundant in Kenosee Lake during the 19<sup>th</sup> century and declined to stable low values until the 2000s, whereas this group varied little over the past 200 yr in White Bear Lake until recent years (Fig. 5). While we currently lack a definitive mechanistic explanation for either the 19<sup>th</sup> century maximum in Kenosee Lake, or the marked difference with nearby White Bear Lake, anecdotally low water levels during the late 19<sup>th</sup> century (Henderson et al. 2002) may have favored anthropogenic eutrophication either from initial settler recreational activities, or cultural use of the lakes by regional First Nations. Although further research is required to resolve the reasons for elevated cyanobacteria during the 19<sup>th</sup> century, the absence of marked increases in *Nostocales* during the 20<sup>th</sup> century contrasts sharply with other lakes that similarly underwent pronounced eutrophication and/or regime shifts (Scheffer et al. 2001; Carpenter and Brock 2006; Bunting et al. 2016).

Preferential increases in eukaryotic phytoplankton (diatoms, cryptophytes, and chlorophytes) over cyanobacteria during the past 200 yr appears to have reduced the mean exposure of phototrophs to UV radiation (Fig. 6). In general, changes in UVR exposure were unrelated to observed lake levels. Significant declines in UVR indices occurred earlier (late 1800s) in Kenosee Lake than in White Bear Lake (~ 1930), similar to timing of significant increases in biomarkers from eukaryotic phytoplankton and declines in lake level (Figs. 2, 5, 6). We infer that reduced UVR exposure arose from progressive, but slightly asynchronous, eutrophication of both lakes, as declines in lake level should have increased mean UVR exposure. Similar declines in UVR exposure are recorded elsewhere in lakes undergoing cultural eutrophication (Stevenson et al. 2016).

Concentrations of most fossil pigments increased significantly after ~ 2000 in both Kenosee and White Bear lakes

(Fig. 5), concomitant with 2–4 m increases in lake level (Fig. 2). In part, these changes reflect post-depositional pigment transformation, as indicated by rapid changes in Chl : pheo ratios in sediments deposited since ~ 2000 (Fig. 6). Such first-order decay is observed in sediments of other prairie lakes (Patoine and Leavitt 2006), but is usually restricted to labile pigments with oxygen- or N-rich functional groups (e.g., Chl *a*, fucoxanthin, and peridinin) rather than less-substituted hydrocarbons (e.g.,  $\beta$ -carotene, alloxanthin, diatoxanthin, lutein-zeaxanthin, etc.; Cuddington and Leavitt 1999; Leavitt and Hodgson 2001). Given that chemically stable pigments also increased markedly toward the surface of the core, we infer that both Kenosee and White Bear lakes may be undergoing modern eutrophication, possibly reflecting increased nutrient transfer from the landscape due to increased runoff and subsequent lake-level rise (McCullough et al. 2012; Tanzeeba and Gan 2012; Asong et al. 2016). Furthermore, elevated concentrations of stable phototrophic biomarkers (e.g., ubiquitous  $\beta$ -carotene) in recent sediments from both study lakes are also consistent with documented water quality problems in both Kenosee and White Bear Lakes since 2000, including the formation of algal blooms (Godwin et al. 2013; Bos et al. 2019).

### Evidence of possible regime shifts in upland lakes

Declines in regional lake level between ~ 1930 and ~ 1950 resulted in the formation of illuminated, anoxic, deep-water habitats that were ideal for the proliferation of obligately anaerobic purple sulfur bacteria (Pfennig 1978) in both Kenosee and White Bear lakes (Fig. 5). The development of such bacterial populations between ~ 1930 and 1950 prerequisites either the illumination of anoxic lake sediments (Jørgensen and Postgate 1982; Maheaux et al. 2016) or the formation of strong seasonal or semi-permanent chemoclines resulting from concentrating solutes associated with climate-mediated lake-level decline (Züllig and Rheineck 1985; Leavitt et al. 1989). We infer that permanent meromictic conditions were not established in either Kenosee Lake or White Bear Lake as there were few concomitant changes in the preservation of labile pigments as okenone concentrations increased (Fig. 6), total okenone concentrations were much lower than those seen in fully meromictic systems (Züllig and Rheineck 1985; Leavitt et al. 1989; Vinebrooke et al. 1998), and peak concentrations of okenone in Kenosee Lake were twofold greater than those in White Bear Lake (Fig. 5), as would be expected given the greater illumination of sediments in the shallower lake following lake-level declines (Table 1; Fig. 6). Alternatively, the observation that both lakes are currently hypersaline (Table 1), despite recent increases in lake level (Fig. 2), suggests that salt concentrations were substantially higher during the lake-level low-stands of the 20<sup>th</sup> century. As shown in other regional lakes, declines in lake level due to evaporative forcing and reduced runoff are associated with higher deep-water salt concentrations that favor establishment

of meromictic conditions in even shallow prairie lakes (Garcés et al. 1995; van der Kamp et al. 2008; Pham et al. 2009).

Alongside increases in pigment concentrations, variance of okenone time series also rose significantly, beginning at ~1930 and reaching maximum values at ~1950 (Fig. 5). Rising variance has been considered as a predictor of a regime shift in some lake systems undergoing eutrophication (Carpenter and Brock 2006; Bunting et al. 2016), although some work suggests that regime shifts can occur without preceding increases in variance, or that rising variance does not always result in a regime shift (Burthe et al. 2016; Ratajczak et al. 2018). As well, it has been noted that rising variance can also be indicative of changes in the variability of environmental forcing agents (e.g., climate, nutrient flux, etc.), independent of the occurrence of regime shifts (Dakos et al. 2015; Burthe et al. 2016). In our case, as pigment variances were calculated from pigment mean concentrations, an increase or decrease in the latter will result in the change in the former, making it difficult to assess the validity of any possible regime changes. Furthermore, differences in temporal averaging of core sediments ( $\text{yr cm}^{-1}$ ) due to physical compression of deposits also makes it difficult to evaluate whether rising variance is truly antecedent to the peak okenone concentrations, despite our attempts to weight samples by temporal resolution in HGAM analyses (see the Methods section). Taken together, these observations suggest that further research is needed to determine how sedimentary records may be used to record changes in variance as a means of distinguishing between abrupt ecosystem changes, regime shifts, or true alternative stable states (Taranu et al. 2018). Regardless, we note that the marked rise in okenone is an unambiguous marker for the establishment of a novel, illuminated, anoxic environment, and that the coeval rise in variance of okenone ca. 1930–1950, but not algal or cyanobacterial pigments, is consistent with potential establishment of an abrupt change in the deep-water environment (Fig. 5). Further research is needed to evaluate these possibilities, including analysis of variance in potential forcing functions, addition of limnological proxies to complement paleolimnological knowledge, and other factors influencing variance (c.f., Bunting et al. 2016).

Periods of illuminated deep-water anoxic habitats appear to have been ephemeral or unstable in both Kenosee and White Bear lakes (Fig. 5). Therefore, in these cases, rising variance does not appear to signal lake transition to an alternate stable state nor represent the establishment of internal feedback mechanisms that favor hysteresis between states (Scheffer et al. 2001; Dakos et al. 2015; Ratajczak et al. 2018). Instead, the >8-m rise and fall of lake levels appears to have resulted in a series of novel phototroph communities whose composition reflected extant hydroclimate and lake-level conditions (McCullough et al. 2012; Asong et al. 2016), physical processes such as mixing and light penetration (Garcés et al. 1995; Hodgson et al. 1998), and landscape influences on nutrient supply (Taranu et al. 2015; Bunting et al. 2016). As well,

despite evidence of warming temperatures and increased nutrient concentrations during the 20<sup>th</sup> century, there was little evidence of increased cyanobacterial populations until after the large concentrations of okenone had abated in the latter half of the century (Fig. 5). These findings are in stark contrast to many studies that highlight increased cyanobacterial production under warm and nutrient-rich conditions such as those present at Kenosee and White Bear lakes (Paerl and Paul 2012; Vogt et al. 2018). We speculate that intervals of reduced mixing, anoxia, or even weak meromixis may have favored internal nutrient supply from sediments that sustained cyanobacteria through the last half of the 20<sup>th</sup> century (Fig. 5). However, we also note that continued lake-level declines may have eventually restricted anoxic, illuminated habitats and sulfur bacterial growth after 1950, possibly due to the influence of high regional winds (Plancq et al. 2018). This hypothesis is consistent with the more pronounced declines in okenone in shallower Kenosee Lake compared to deeper White Bear Lake (Fig. 5).

## Conclusions

Kenosee and White Bear lakes have experienced >8 m declines in lake level over the last century due to increased temperature and evaporation rates, combined with variations in climate systems that regulate introduction of moist oceanic air and precipitation (Akinremi et al. 1999; Bonsal et al. 2006; Michels et al. 2007; Pham et al. 2009). Quantification of historical changes in sedimentary stable isotopes, geochemistry, and biomarker pigments suggests that both lakes began to eutrophy during the early 20<sup>th</sup> century, coeval with documented declines in lake levels (Fig. 2). By the mid-20<sup>th</sup> century, declines in lake levels and increased solute concentrations allowed the development of illuminated anoxic habitats that allowed expansion of obligately anaerobic purple sulfur bacteria (Züllig and Rheineck 1985; Leavitt et al. 1989). However, despite evidence of eutrophication in biomarkers from eukaryotic phytoplankton, Kenosee and White Bear lakes did not exhibit substantial increases in colonial cyanobacteria, nor the significant increases in temporal variation of primary producers typical of lakes undergoing nutrient-driven regime shifts (Taranu et al. 2015, 2018; Bunting et al. 2016). Instead, the establishment of anoxic habitats for photosynthetic sulfur bacteria was recorded by rising variance only in their biomarker okenone, and suggests that only deep-water habitat and taxa, rather than the entire ecosystem, underwent a transition to a novel structure. Overall, it appears that rising variance of okenone did not constitute either an alternate state change or true regime shift (sensu Burthe et al. 2016; Ratajczak et al. 2018), but instead illustrates that climate-induced changes in lake-level can result in abrupt variation in lake structure that reconfigures deep-water habitats and biological communities (Fig. 5). Given that GCMs predict that the northern Great Plains region will rapidly become more arid in

coming century (Tanzeeba and Gan 2012; Asong et al. 2016; Sauchyn et al. 2020), we anticipate the development of similar deep-water and biotic assemblages in other regional lakes.

### Data Availability Statement

Data from this study is open and available on github at <https://github.com/simpson-lab/kenosee-white-bear>.

### References

- Agriculture and Agri-Food Canada. 2013. Annual crop inventory. Ottawa, Ontario, Canada. <https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9>
- Akinremi, O. O., S. M. McGinn, and H. W. Cutforth. 1999. Precipitation trends on the Canadian prairies. *J. Climate* **12**: 2996–3003.
- Appelby, P. G., P. J. Nolan, D. W. Gifford, M. J. Godfrey, F. Oldfield, N. J. Anderson, and R. W. Battarbee. 1986.  $^{210}\text{Pb}$  dating by low background gamma counting. *Hydrobiologia* **143**: 21–17. doi:[10.1007/BF00026640](https://doi.org/10.1007/BF00026640)
- Asong, Z. E., M. N. Khalil, and H. S. Wheater. 2016. Projected changes in precipitation and temperature over the Canadian Prairie Provinces using the Generalized Linear Model statistical downscaling approach. *J. Hydrol.* **539**: 429–446. doi:[10.1016/j.jhydrol.2016.05.044](https://doi.org/10.1016/j.jhydrol.2016.05.044)
- Binford, M. W. 1990. Calculation and uncertainty analysis of  $^{210}\text{Pb}$  dates for PIRLA project lake sediment cores. *J. Paleolimnol.* **3**: 253–267. doi:[10.1007/BF00219461](https://doi.org/10.1007/BF00219461)
- Bonsal, B. R., T. D. Prowse, C. R. Duguay, and M. P. Lacroix. 2006. Impacts of large-scale teleconnections on freshwater-ice break/freeze-up dates over Canada. *J. Hydrol.* **330**: 340–353. doi:[10.1016/j.jhydrol.2006.03.022](https://doi.org/10.1016/j.jhydrol.2006.03.022)
- Bos, J. S., L. Nanayakkara, M. Hurlbert, and K. Finlay. 2019. Citizen science for Saskatchewan lakes: A pilot project. *Lake Reserv. Manag.* **35**: 77–89. doi:[10.1080/10402381.2018.1538172](https://doi.org/10.1080/10402381.2018.1538172)
- Botrel, M., I. Gregory-Eaves, and R. Maranger. 2014. Defining drivers of nitrogen stable isotopes ( $\delta^{15}\text{N}$ ) of surface sediments in temperate lakes. *J. Paleolimnol.* **52**: 419–433. doi:[10.1007/s10933-014-9802-6](https://doi.org/10.1007/s10933-014-9802-6)
- Bryson, R. A., and F. K. Hare. 1974. Climates of North America. Elsevier.
- Bunting, L., P. R. Leavitt, G. L. Simpson, B. Wissel, K. R. Laird, B. F. Cumming, A. St. Amand, and D. R. Engstrom. 2016. Increased variability and sudden ecosystem state change in Lake Winnipeg, Canada, caused by 20th century agriculture. *Limnol. Oceanogr.* **61**: 2090–2107. doi:[10.1002/limo.10355](https://doi.org/10.1002/limo.10355)
- Burthe, S. J., and others. 2016. Do early warning indicators consistently predict nonlinear change in long-term ecological data? *J. Appl. Ecol.* **53**: 666–676.
- Carpenter, S. R., and W. A. Brock. 2006. Rising variance: A leading indicator of ecological transition. *Ecol. Lett.* **9**: 311–318. doi:[10.1111/j.1461-0248.2005.00877.x](https://doi.org/10.1111/j.1461-0248.2005.00877.x)
- Carpenter, S. R., E. H. Stanley, and M. J. Vander Zanden. 2011. State of the world's freshwater ecosystems: Physical, chemical, and biological changes. *Annu. Rev. Environ. Res.* **36**: 75–99. doi:[10.1146/annurev-environ-021810-094524](https://doi.org/10.1146/annurev-environ-021810-094524)
- Coles, A. E., and J. J. McDonnell. 2018. Fill and spill drives runoff connectivity over frozen ground. *J. Hydrol.* **558**: 115–128. doi:[10.1016/j.jhydrol.2018.01.016](https://doi.org/10.1016/j.jhydrol.2018.01.016)
- Cottingham, K. L., J. A. Rusak, and P. R. Leavitt. 2000. Increased ecosystem variability and reduced predictability following fertilization: Evidence from palaeolimnology. *Ecol. Lett.* **3**: 340–348. doi:[10.1046/j.1461-0248.2000.00158.x](https://doi.org/10.1046/j.1461-0248.2000.00158.x)
- Cuddington, K., and P. R. Leavitt. 1999. An individual-based model of pigment flux in lakes: Implications for organic biogeochemistry and paleoecology. *Can. J. Fish. Aquat. Sci.* **56**: 1964–1977. doi:[10.1139/f99-108](https://doi.org/10.1139/f99-108)
- Cullimore, D. R., and D. Griffin. 1979. Data collection relevant to White Bear Lake water quality status, 24. Regina Water Research Institute, Univ. of Regina.
- Dakos, V., S. R. Carpenter, E. H. van Nes, and M. Scheffer. 2015. Resilience indicators: Prospects and limitations of early warnings of regime shifts. *Phil. Trans. R. Soc. B* **37**: 20130263. doi:[10.1098/rstb.2013.0263](https://doi.org/10.1098/rstb.2013.0263)
- Davis, T., D. Berry, G. Boyer, and C. Gobler. 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. *Harmful Algae* **8**: 715–725. doi:[10.1016/j.hal.2009.02.004](https://doi.org/10.1016/j.hal.2009.02.004)
- Dibike, Y., T. Prowse, B. Bonsal, and H. O'Neil. 2016. Implications of future climate on water availability in the western Canadian river basins. *Int. J. Climatol.* **37**: 3247–3263. doi:[10.1002/joc.4912](https://doi.org/10.1002/joc.4912)
- Dröscher, I., A. Patoine, K. Finlay, and P. R. Leavitt. 2009. Climate control of spring clear-water phase through the transfer of energy and mass to lakes. *Limnol. Oceanogr.* **54**: 2469–2480. doi:[10.4319/lo.2009.54.6\\_part\\_2.2469](https://doi.org/10.4319/lo.2009.54.6_part_2.2469)
- Dumanski, S., J. W. Pomeroy, and C. J. Westbrook. 2015. Hydrological regime changes in a Canadian Prairie basin. *Hydrol. Process.* **29**: 3893–3904. doi:[10.1002/hyp.10567](https://doi.org/10.1002/hyp.10567)
- Fang, X., A. Minke, J. Pomeroy, T. Brown, C. Westbrook, X. Guo, and S. Guangul. 2007. A review of Canadian Prairie hydrology: Principles, modelling and response to land use and drainage change. Center for Hydrology Report, 2. Univ. of Saskatchewan.
- Fischer, P., and U. Öhl. 2005. Effects of water-level fluctuations on the littoral benthic fish community in lakes: A mesocosm experiment. *Behav. Ecol.* **16**: 741–746. doi:[10.1093/beheco/ari047](https://doi.org/10.1093/beheco/ari047)
- Fritz, S. C. 1990. Twentieth-century salinity and water-level fluctuations in Devils Lake, North Dakota: Test of a diatom-based transfer function. *Limnol. Oceanogr.* **35**: 1771–1781. doi:[10.4319/lo.1990.35.8.1771](https://doi.org/10.4319/lo.1990.35.8.1771)
- Gaeta, J. W., G. G. Sass, and S. R. Carpenter. 2014. Drought-driven lake level decline: Effects on coarse woody habitat

- and fishes. *Can. J. Fish. Aquat. Sci.* **71**: 315–325. doi:[10.1139/cjfas-2013-0451](https://doi.org/10.1139/cjfas-2013-0451)
- Gal, G., M. Skerjanec, and N. Atanasova. 2013. Fluctuations in water level and the dynamics of zooplankton: A data-driven modelling approach. *Freshw. Biol.* **58**: 800–816. doi:[10.1111/fwb.12087](https://doi.org/10.1111/fwb.12087)
- Garcés, B. L. V., K. Kelts, and E. Ito. 1995. Oxygen and carbon isotope trends and sedimentological evolution of a meromictic and saline lacustrine system: The Holocene Medicine Lake basin, North American Great Plains, USA. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **117**: 253–278. doi:[10.1016/0031-0182\(94\)00136-V](https://doi.org/10.1016/0031-0182(94)00136-V)
- Glew, J. R. 1989. A new trigger mechanism for sediment samplers. *J. Paleolimnol.* **2**: 241–243.
- Godwin, B., V. Wittrock, and J. Thorpe. 2013. Ecosystem based surface water quantity management plan for Moose Mountain Provincial Park. Saskatchewan Research Council Publication No. 13413-1E13, Regina, Saskatchewan.
- Gu, B., A. D. Chapman, and C. L. Schelske. 2006. Factors controlling seasonal variations in stable isotope composition of particulate organic matter in a softwater eutrophic lake. *Limnol. Oceanogr.* **51**: 2837–2848. doi:[10.4319/lo.2006.51.6.2837](https://doi.org/10.4319/lo.2006.51.6.2837)
- Haig, H. A., N. M. Hayes, G. L. Simpson, Y. Yi, B. Wissel, K. Finlay, K. R. Hodder, and P. R. Leavitt. 2021. Application of isotope mass balance models to assess spatial variability in lake hydrology on the Northern Great Plains, Canada. *J. Hydrol. X In review.* **6**: 100046.
- Haig, H. A., N. M. Hayes, G. L. Simpson, Y. Yi, B. Wissel, K. R. Hodder, and P. R. Leavitt. 2020. Comparison of isotopic mass balance and instrumental techniques as estimates of basin hydrology in seven connected lakes over 12 years. *J. Hydrol. X* **6**: 100046. doi:[10.1016/j.hydroa.2019.100046](https://doi.org/10.1016/j.hydroa.2019.100046)
- Hambright, K. D., T. Zohary, W. Eckert, S. S. Schwartz, C. L. Schelske, K. R., and P. R. Leavitt PR. 2008. Exploitation and destabilization of a warm, freshwater ecosystem through engineered hydrological change. *Ecol. App.* **18**: 1591–1603. doi:[10.1890/07-0931.1](https://doi.org/10.1890/07-0931.1),
- Hayes, N. M., A. Patoine, H. A. Haig, G. L. Simpson, V. J. Swarbrick, E. Wiik, and P. R. Leavitt. 2019. Spatial and temporal variation in nitrogen fixation and its importance to phytoplankton growth in phosphorus-rich lakes. *Freshw. Biol.* **64**: 269–283. doi:[10.1111/fwb.13214](https://doi.org/10.1111/fwb.13214)
- Henderson, N., E. Hogg, E. Barrow, and B. Dolter. 2002. Climate change impacts on the island forests of the Great Plains and the implications for nature conservation policy: The outlook for Sweet Grass Hills (Montana), Cypress Hills (Alberta-Saskatchewan), Moose Mountain (Saskatchewan), Spruce Woods (Manitoba) and Turtle Mountain (Manitoba-North Dakota). Prairie Adaptation and Research Collaborative, Univ. of Regina.
- Hodgson, D. A., S. W. Wright, P. A. Tyler, and N. Davies. 1998. Analysis of fossil pigments from algae and bacteria in meromictic Lake Fidler, Tasmania, and its application to lake management. *J. Paleolimnol.* **19**: 1–22. doi:[10.1023/A:1007909018527](https://doi.org/10.1023/A:1007909018527)
- Huisman, J., G. A. Codd, H. W. Paerl, B. W. Ibelings, J. M. Verspagen, and P. M. Visser. 2018. Cyanobacterial blooms. *Nat. Rev. Microbiol.* **16**: 471–483. doi:[10.1038/s41579-018-0040-1](https://doi.org/10.1038/s41579-018-0040-1)
- Hurrell, J. W. 1995. Decadal trends in the North Atlantic oscillation: Regional temperatures and precipitation. *Science* **269**: 676–679. doi:[10.1126/science.269.5224.676](https://doi.org/10.1126/science.269.5224.676)
- Jeppesen, E., and others. 2014. Climate change impacts on lakes: An integrated ecological perspective based on a multi-faceted approach, with special focus on shallow lakes. *J. Limnol.* **73**: 84–107. doi:[10.4081/jlimnol.2014.844](https://doi.org/10.4081/jlimnol.2014.844)
- Jørgensen, B. B., and J. R. Postgate. 1982. Ecology of the bacteria of the sulphur cycle with special reference to anoxic-oxic interface environments. *Philos. Trans. R. Soc. Lond. B* **298**: 543–561. doi:[10.1098/rstb.1982.0096](https://doi.org/10.1098/rstb.1982.0096)
- van der Kamp, G., D. Keir, and M. Evans. 2008. Long-term water level changes in closed-basin lakes of the Canadian prairies. *Can. Water Resour. J.* **33**: 23–38. doi:[10.4296/cwrj3301023](https://doi.org/10.4296/cwrj3301023)
- van der Kamp, G., and H. Maathuis. 1991. Annual fluctuations of groundwater levels as a result of loading by surface moisture. *J. Hydrol.* **127**: 137–152. doi:[10.1016/0022-1694\(91\)90112-U](https://doi.org/10.1016/0022-1694(91)90112-U)
- Leavitt, P. R., R. Brock, C. Ebel, and A. Patoine. 2006. Landscape-scale effects of urban nitrogen on a chain of freshwater lakes in central North America. *Limnol. Oceanogr.* **51**: 2262–2277. doi:[10.4319/lo.2006.51.5.2262](https://doi.org/10.4319/lo.2006.51.5.2262)
- Leavitt, P. R., S. R. Carpenter, and J. F. Kitchell. 1989. Whole-lake experiments: The annual record of fossil pigments and zooplankton. *Limnol. Oceanogr.* **34**: 700–717. doi:[10.4319/lo.1989.34.4.0700](https://doi.org/10.4319/lo.1989.34.4.0700)
- Leavitt, P. R., and D. A. Hodgson. 2001. Sedimentary pigments, p. 95–325. In J. P. Smol, H. J. B. Birks, and W. M. Last [eds.], *Tracking environmental change using lake sediments, Volume 3: Terrestrial, algal, and siliceous indicators*. Kluwer Academic Publishers.
- Leavitt, P. R., D. E. Schindler, A. J. Paul, A. K. Hardie, and D. W. Schindler. 1994. Fossil pigment records of phytoplankton in trout-stocked alpine lakes. *Can. J. Fish. Aquat. Sci.* **51**: 2411–2423. doi:[10.1139/f94-241](https://doi.org/10.1139/f94-241)
- Leavitt, P. R., R. D. Vinebrooke, D. Donald, J. P. Smol, and D. W. Schindler. 1997. Past ultraviolet radiation environments in lakes derived from fossil pigments. *Nature* **388**: 457–459. doi:[10.1038/41296](https://doi.org/10.1038/41296)
- Liu, J., R. E. Stewart, and K. Szeto. 2008. Water vapor fluxes over the Canadian Prairies and the Mackenzie River Basin, p. 109–125. In M. Woo [ed.], *Cold region atmospheric and hydrologic studies*. Springer.
- Madigan, M. T. 1995. Microbiology of nitrogen fixation by anoxygenic photosynthetic bacteria, p. 915–928. In R. E. Blankenship, M. T. Madigan, and C. T. Bauer [eds.], *Anoxygenic photosynthetic bacteria. Advances in*

- photosynthesis and respiration book series (AIPH, volume 2). Springer.
- Maheaux, H., P. R. Leavitt, and L. Jackson. 2016. Asynchronous onset of eutrophication among shallow prairie lakes of the Northern Great Plains, Alberta, Canada. *Glob. Change Biol.* **22**: 271–283. doi:[10.1111/gcb.13076](https://doi.org/10.1111/gcb.13076)
- Mantua, N., S. Hare, Y. Zhang, J. W. Wallace, and W. R. Francis. 1997. A pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Am. Meteorol. Soc.* **78**: 1069–1079. doi:[10.1175/1520-0477\(1997\)078%3C1069:APICOW%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078%3C1069:APICOW%3E2.0.CO;2)
- Mao, D., and K. A. Cherkauer. 2008. Impacts of land-use change on hydrologic responses in the Great Lakes region. *J. Hydrol.* **374**: 71–82. doi:[10.1016/j.jhydrol.2009.06.016](https://doi.org/10.1016/j.jhydrol.2009.06.016)
- McCullough, G. K., S. B. Page, R. H. Hesslein, M. P. Stanton, H. J. Kling, A. G. Salki, and D. J. Barbera. 2012. Hydrological forcing of a recent trophic surge in Lake Winnipeg. *J. Great Lakes Res.* **38**: 95–105. doi:[10.1016/j.jglr.2011.12.012](https://doi.org/10.1016/j.jglr.2011.12.012)
- Meyers, P. A., and J. L. Teranes. 2001. Sediment organic matter, p. 239–270. In W. M. Last and J. P. Smol [eds.], *Tracking environmental change using Lake sediments, volume 2: Physical and geochemical methods*. Kluwer Academic Publishers.
- Michels, A., K. R. Laird, S. E. Wilson, D. Thomson, P. R. Leavitt, R. J. Oglesby, and B. F. Cumming. 2007. Multi-decadal to millennial-scale shifts in drought conditions on the Canadian Prairies over the past six millennia: Implications for future drought assessment. *Glob. Change Biol.* **13**: 1295–1307. doi:[10.1111/j.1365-2486.2007.01367.x](https://doi.org/10.1111/j.1365-2486.2007.01367.x)
- Oduor, S. O., and M. Schagerl. 2007. Temporal trends of ion contents and nutrients in three Kenyan Rift Valley saline-alkaline lakes and their influence on phytoplankton biomass, p. 59–68. In R. D. Gulati, E. Lammens, N. de Pauw, and E. van Donk [eds.], *Shallow lakes in a changing world*. Springer.
- Paerl, H. W., and V. J. Paul. 2012. Climate change: Links to global expansion of harmful cyanobacterial blooms. *Nat. Rev. Microbiol.* **16**: 471–483. doi:[10.1016/j.nrmicro.2011.08.002](https://doi.org/10.1016/j.nrmicro.2011.08.002)
- Patoine, A., and P. R. Leavitt. 2006. Century-long synchrony of algal fossil pigments in a chain of Canadian prairie lakes. *Ecology* **87**: 1710–1721. doi:[10.1890/0012-9658](https://doi.org/10.1890/0012-9658)
- Pedersen, E. J., D. L. Miller, G. L. Simpson, and N. Ross. 2019. Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **7**: e6876. doi:[10.7717/peerj.6876](https://doi.org/10.7717/peerj.6876)
- Pfennig, N. 1978. General physiology and ecology of photosynthetic bacteria, p. 3–18. In R. K. Clayton and W. R. Sistrom [eds.], *The photosynthetic bacteria*. Plenum Press.
- Pham, S. V., P. R. Leavitt, S. McGowan, and P. Peres-Neto. 2008. Spatial variability of climate and land-use effects on lakes of the northern Great Plains. *Limnol. Oceanogr.* **53**: 728–742. doi:[10.4319/lo.2008.53.2.0728](https://doi.org/10.4319/lo.2008.53.2.0728)
- Pham, S. V., P. R. Leavitt, S. McGowan, B. Wissel, and L. Wassenaar. 2009. Spatial and temporal variability of prairie lake hydrology as revealed using stable isotopes of hydrogen and oxygen. *Limnol. Oceanogr.* **54**: 101–118. doi:[10.4319/lo.2009.54.1.0101](https://doi.org/10.4319/lo.2009.54.1.0101)
- Plancq, J., B. Cavazzin, S. Juggins, H. A. Haig, P. R. Leavitt, and J. L. Toney. 2018. Assessing environmental controls on the distribution of long-chain alkenones in the Canadian prairies. *Org. Geochem.* **117**: 43–55. doi:[10.1016/j.orggeochem.2017.12.005](https://doi.org/10.1016/j.orggeochem.2017.12.005)
- Pomeroy, J., D. Gray, T. Brown, N. Hedstrom, W. Quinton, R. Granger, and S. Carey. 2007. The cold regions hydrological model: A platform for basing process representation and model structure on physical evidence. *Hydrolog. Process.* **21**: 2650–2667. doi:[10.1002/hyp.6787](https://doi.org/10.1002/hyp.6787)
- Pya, N. 2021. SCAM: Shape constrained additive models. <https://cran.r-project.org/package=scam>
- R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Randsalu-Wendrup, L., D. J. Conley, J. Carstensen, and S. C. Fritz. 2016. Paleolimnological records of regime shifts in lakes in response to climate change and anthropogenic activities. *J. Paleolimnol.* **56**: 1–14. doi:[10.1007/s10933-016-9884-4](https://doi.org/10.1007/s10933-016-9884-4)
- Ratajczak, Z., and others. 2018. Abrupt change in ecological systems: Inference and diagnosis. *Trends Ecol. Evol.* **33**: 513–526. doi:[10.1016/j.tree.2018.04.013](https://doi.org/10.1016/j.tree.2018.04.013)
- Sauchyn, D. J., D. Davidson, and M. Johnston. 2020. Prairie Provinces, chapter 4. In F. J. Warren, N. Lulham, and D. S. Lemmen [eds.], *Canada in a changing climate: Regional perspectives report*. Government of Canada.
- Savage, C., P. R. Leavitt, and R. Elmgren. 2004. Distribution and retention of effluent nitrogen in surface sediments of a coastal bay. *Limnol. Oceanogr.* **49**: 1503–1511. doi:[10.4319/lo.2004.49.5.1503](https://doi.org/10.4319/lo.2004.49.5.1503)
- Scheffer, M., and S. R. Carpenter. 2003. Catastrophic regime shifts in ecosystems: Linking theory to observation. *Trends Ecol. Evol.* **18**: 648–656. doi:[10.1016/j.tree.2003.09.002](https://doi.org/10.1016/j.tree.2003.09.002)
- Scheffer, M., S. R. Carpenter, J. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. *Nature* **413**: 591–596. doi:[10.1038/35098000](https://doi.org/10.1038/35098000)
- Sereda, J., M. Bogard, J. Hudson, D. Helps, and T. Dessouki. 2011. Climate warming and the onset of salinization: Rapid changes in the limnology of two northern plains lakes. *Limnologica* **41**: 1–9. doi:[10.1016/j.limno.2010.03.002](https://doi.org/10.1016/j.limno.2010.03.002)
- Shaw, R. D., and E. E. Prepas. 1990. Groundwater-lake interactions: II. Nearshore seepage patterns and the contributions of groundwater to lakes in central Alberta. *J. Hydrol.* **119**: 121–136. doi:[10.1016/0022-1694\(90\)90038-Y](https://doi.org/10.1016/0022-1694(90)90038-Y)
- Simpson, G. L. 2018. Modelling palaeoecological time series using generalised additive models. *Front. Ecol. Evol.* **6**: 149. doi:[10.3389/fevo.2018.00149](https://doi.org/10.3389/fevo.2018.00149)

- Simpson, G. L. 2021. Gratia: Graceful ‘ggplot’-based graphics and other functions for GAMs fitted using ‘mgcv’. <https://cran.r-project.org/package=gratia>.
- Søndergaard, M., R. Bjerring, and E. Jeppesen. 2013. Persistent internal phosphorus loading during summer in shallow eutrophic lakes. *Hydrobiologia* **710**: 95–107. doi:[10.1007/s10750-012-1091-3](https://doi.org/10.1007/s10750-012-1091-3)
- Stevenson, M. A., S. McGowan, N. J. Anderson, R. H. Foy, P. R. Leavitt, Y. R. McElearney, D. R. Engstrom, and S. Pla-Rabés. 2016. Impacts of forest plantation management on primary production in upland lakes from north-west Ireland. *Glob. Chang. Biol.* **22**: 1490–1504. doi:[10.1111/gcb.13194](https://doi.org/10.1111/gcb.13194)
- Tanzeeba, S., and T. Gan. 2012. Potential impact of climate change on the water availability of South Saskatchewan River Basin. *Clim. Change* **112**: 355–386. doi:[10.1007/s10584-011-0221-7](https://doi.org/10.1007/s10584-011-0221-7)
- Taranu, Z. E., S. R. Carpenter, M. E. Perga, T. Ives, V. Frossard, Z. Thoma, J. P. Jenny, and J. C. Vermaire. 2018. Can we detect ecosystem critical transitions and early warning signals of catastrophic shifts from paleo-ecological records? *Ecosphere* **9**: e02438. doi:[10.1002/ecs2.2438](https://doi.org/10.1002/ecs2.2438)
- Taranu, Z. E., and others. 2015. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol. Lett.* **18**: 375–384. doi:[10.1111/ele.12420](https://doi.org/10.1111/ele.12420)
- Trenberth, K. E., and J. W. Hurrell. 1994. Decadal atmosphere-ocean variations in the Pacific. *Clim. Dyn.* **9**: 303–319. doi:[10.1007/BF00204745](https://doi.org/10.1007/BF00204745)
- Vance, R. E., W. M. Last, and A. J. Smith. 1997. Hydrologic and climatic implications of a multidisciplinary study of late Holocene sediment from Kenosee Lake, southeastern Saskatchewan, Canada. *J. Paleolimnol.* **18**: 365–393. doi:[10.1023/A:1007948909342](https://doi.org/10.1023/A:1007948909342)
- Vinebrooke, R., R. Hall, P. R. Leavitt, and B. F. Cumming. 1998. Fossil pigments as indicators of phototrophic response to salinity and climatic change in lakes of western Canada. *Can. J. Fish. Aquat. Sci.* **55**: 668–681. doi:[10.1139/cjfas-55-3-668](https://doi.org/10.1139/cjfas-55-3-668)
- Vogt, R. J., S. Sharma, and P. R. Leavitt. 2018. Direct and interactive effects of climate, meteorology, river hydrology, and lake characteristics on water quality in productive lakes of the Canadian prairies. *Can. J. Fish. Aquat. Sci.* **75**: 47–59. doi:[10.1139/cjfas-2016-0520](https://doi.org/10.1139/cjfas-2016-0520)
- Vörösmarty, C. J., P. Green, J. Salisbury, and R. B. Lammers. 2000. Global water resources: Vulnerability from climate change and population growth. *Science* **289**: 284–288. doi:[10.1126/science.289.5477.284](https://doi.org/10.1126/science.289.5477.284)
- Wickham, H. 2021. Tidyr: Tidy messy data. <https://cran.r-project.org/package=tidyr>
- Wickham, H., R. François, L. Henry, and K. Müller. 2021. Dplyr: A grammar of data manipulation. <https://cran.r-project.org/package=dplyr>
- Wickham, H., and others. 2016. ggplot2: Elegant graphics for data analysis. <https://cran.r-project.org/package=ggplot2>
- Wigdahl, C. R., J. E. Saros, S. C. Fritz, J. R. Stone, and D. R. Engstrom. 2014. The influence of basin morphometry on the regional coherence of patterns of diatom-inferred salinity in lakes of the northern Great Plains (USA). *Holocene* **24**: 603–613. doi:[10.1177/0959683614523154](https://doi.org/10.1177/0959683614523154)
- Wilke, C. O. 2020. Cowplot: Streamlined plot theme and plot annotations for ‘ggplot2’. <https://cran.r-project.org/package=cowplot>
- Wissel, B., R. N. Cooper, P. R. Leavitt, and S. V. Pham. 2011. Hierarchical regulation of pelagic invertebrates in lakes of the northern Great Plains: A novel model for inter-decadal effects of future climate change. *Glob. Change Biol.* **17**: 172–185. doi:[10.1111/j.1365-2486.2010.02291.x](https://doi.org/10.1111/j.1365-2486.2010.02291.x)
- Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. B* **73**: 3–36. doi:[10.1111/j.1467-9868.2010.00749.x](https://doi.org/10.1111/j.1467-9868.2010.00749.x)
- Wood, S. N. 2017. Generalized additive models: An introduction with R, 2nd ed. Chapman Hall.
- Woodward, C. A., A. P. Potito, and D. W. Beilman. 2012. Carbon and nitrogen stable isotope ratios in surface sediments from lakes of western Ireland: Implications for inferring past lake productivity and nitrogen loading. *J. Paleolimnol.* **47**: 167–184. doi:[10.1007/s10933-011-9568-z](https://doi.org/10.1007/s10933-011-9568-z)
- Xiao, K., and others. 2018. Evaporation from a temperate closed-basin lake and its impact on present, past, and future water level. *J. Hydrol.* **561**: 59–75. doi:[10.1016/j.jhydrol.2018.03.059](https://doi.org/10.1016/j.jhydrol.2018.03.059)
- Zhou, X., G. Huang, X. Wang, and G. Cheng. 2018. Dynamically-downscaled temperature and precipitation changes over Saskatchewan using the PRECIS model. *Clim. Dyn.* **50**: 1321–1334. doi:[10.1007/s00382-017-3687-9](https://doi.org/10.1007/s00382-017-3687-9)
- Zinabu, G. M. 2002. The effects of wet and dry seasons on concentrations of solutes and phytoplankton biomass in seven Ethiopian rift-valley lakes. *Limnologica* **32**: 169–179. doi:[10.1016/S0075-9511\(02\)80006-8](https://doi.org/10.1016/S0075-9511(02)80006-8)
- Zohary, T., and I. Ostrovsky. 2011. Ecological impacts of excessive water level fluctuations in stratified freshwater lakes. *Inland Waters* **1**: 47–59. doi:[10.5268/IW-1.1.406](https://doi.org/10.5268/IW-1.1.406)
- Züllig, H., and S. Rheineck. 1985. Pigmente phototropher Bakterien in Seesedimenten und ihre Bedeutung für die Seeforschung. *Swiss J. Hydrol.* **47**: 87–126. doi:[10.1007/BF02551937](https://doi.org/10.1007/BF02551937)

### Acknowledgments

The authors thank Kyle Hodder for GIS analysis using Arc 10.4.1 and ArcGISPro-2.0.1, Bjoern Wissel for stable isotope and geochemical analyses, and Zoraida Quinones-Rivera for assistance with HPLC analysis. We also thank Tiffany Blampied, Jessica Bos, and Corey McCowan for assistance with field sampling. Funding for this project was provided through NSERC Discovery grants to G.L.S., P.R.L., and K.F., the Saskatchewan Ministry of Parks, Culture, and Sport to K.F., the Canada Foundation for Innovation, the Canada Research Chair program, the Province of Saskatchewan, and University of Regina. We also thank two anonymous reviewers for their helpful comments that greatly improved the quality of this manuscript. Data in this study were obtained from samples collected

from part of Canada's Treaty 4, the traditional territories of Nêhiyawak, Anihšināpēk, Dakota, Lakota, and Nakoda First Nations, and the home-land of the Métis Nation.

**Conflict of Interest**

None declared.

Submitted 17 August 2021

Revised 02 February 2022

Accepted 20 February 2022

Associate editor: Grace M Wilkinson

RESEARCH

Open Access



# Movement ecology of vulnerable lowland tapirs between areas of varying human disturbance

E. P. Medici<sup>1,2,3\*</sup>, S. Mezzini<sup>4</sup>, C. H. Fleming<sup>5,6</sup>, J. M. Calabrese<sup>7,8,9</sup> and M. J. Noonan<sup>4</sup>

## Abstract

**Background:** Animal movement is a key ecological process that is tightly coupled to local environmental conditions. While agriculture, urbanisation, and transportation infrastructure are critical to human socio-economic improvement, these have spurred substantial changes in animal movement across the globe with potential impacts on fitness and survival. Notably, however, human disturbance can have differential effects across species, and responses to human activities are thus largely taxa and context specific. As human disturbance is only expected to worsen over the next decade it is critical to better understand how species respond to human disturbance in order to develop effective, case-specific conservation strategies.

**Methods:** Here, we use an extensive telemetry dataset collected over 22 years to fill a critical knowledge gap in the movement ecology of lowland tapirs (*Tapirus terrestris*) across areas of varying human disturbance within three biomes in southern Brazil: the Pantanal, Cerrado, and Atlantic Forest.

**Results:** From these data we found that the mean home range size across all monitored tapirs was 8.31 km<sup>2</sup> (95% CI 6.53–10.42), with no evidence that home range sizes differed between sexes nor age groups. Interestingly, although the Atlantic Forest, Cerrado, and Pantanal vary substantially in habitat composition, levels of human disturbance, and tapir population densities, we found that lowland tapir movement behaviour and space use were consistent across all three biomes. Human disturbance also had no detectable effect on lowland tapir movement. Lowland tapirs living in the most altered habitats we monitored exhibited movement behaviour that was comparable to that of tapirs living in a near pristine environment.

**Conclusions:** Contrary to our expectations, although we observed individual variability in lowland tapir space use and movement, human impacts on the landscape also had no measurable effect on their movement. Lowland tapir movement behaviour thus appears to exhibit very little phenotypic plasticity in response to human disturbance. Crucially, the lack of any detectable response to anthropogenic disturbance suggests that human modified habitats risk being ecological traps for tapirs and this information should be factored into conservation actions and species management aimed towards protecting lowland tapir populations.

**Keywords:** Anthropocene, Continuous-time movement modelling, Home range, Human Footprint Index, Space use

## Background

While agriculture, urbanisation, and transportation infrastructure are critical to human socio-economic improvement [17], the associated habitat transformations represent a major threat to species survival [18, 54, 76]. Of particular concern is the impact of human activities

\*Correspondence: medici@ipe.org.br

<sup>1</sup> Lowland Tapir Conservation Initiative (LTCI), Instituto de Pesquisas Ecológicas (IPÊ), Rodovia Dom Pedro I, km 47, Nazaré Paulista, São Paulo 12960-000, Brazil

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

on animal movement and space use [3, 16, 71]. Animal movement governs how individuals, populations, and species interact with each other and the environment [29, 39, 63] and mediates key ecological processes [6]. The capacity for individuals to move unhindered across complex landscapes is therefore critical for species survival and ecosystem function. Problematically, human development has reduced the amount of habitat available to wildlife [8, 13, 32]. This has spurred substantial changes in animal movement behaviour across the globe [16, 19, 71], with potential consequences including reduced fitness and survival, altered predator–prey dynamics, reduced seed dispersal, genetic isolation and local extinction [14, 15, 19, 72].

Notably, human disturbance has been shown to have differential effects across species [16, 70], even for closely related taxa occupying the same habitat [68]. Responses to human activities are thus largely context specific [16] and cannot be expected to be consistent across taxa. For instance, while Wall et al. [78] found a tendency for African elephants (*Loxodonta spp.*) to exhibit reduced movement in human modified landscapes, Morato et al. [47] noted that jaguars (*Panthera onca*) living in regions with high human population densities in South America occupied home ranges that were orders of magnitude larger than those of jaguars living in more pristine habitats. As human disturbance is only expected to worsen over the next decade it is critical to better understand how species respond to human disturbance to develop effective, case-specific conservation strategies.

Here we focus on understanding how the movement behaviour of lowland tapirs (*Tapirus terrestris*) varies across areas of varying human disturbance within the Pantanal, Cerrado, and Atlantic Forest biomes in southern Brazil. Lowland tapirs are herbivores of the order Perissodactyla that can reach over 2.5 m in length and weigh up to 250 kg [40]. While lowland tapirs are distributed throughout South America [27], their populations have suffered severe reductions, with local and regional extirpations, and are currently classified as vulnerable to extinction [74]. Although the incorporation of information on animal movement is a key component in designing effective conservation and recovery strategies [3], currently, very little is known about the movement ecology of tapirs (but see [22, 52, 69]). This knowledge gap is especially pertinent given that large terrestrial mammals, such as tapirs, tend to have larger home ranges and greater absolute mobility than do small mammals [11, 51], making them more susceptible to anthropogenic impacts than smaller bodied species [31, 71]. Here, we use an extensive telemetry dataset collected over 22 years to describe the movement ecology of tapirs and study how changes in habitat composition and human

disturbance influence their movement and space use. First, animals living in highly productive environments do not need to range over wide areas to meet their energetic needs [35, 48, 57]. As such, we expected that tapirs should exhibit plasticity in their movement and space use in relation to local environmental conditions as well as biome type. Furthermore, because human activity tends to result in increased movement for large herbivores [16] our underlying hypothesis was that tapirs should exhibit greater movement distances and larger home-range areas when living in human-modified landscapes.

## Methods

### Study area and data collection

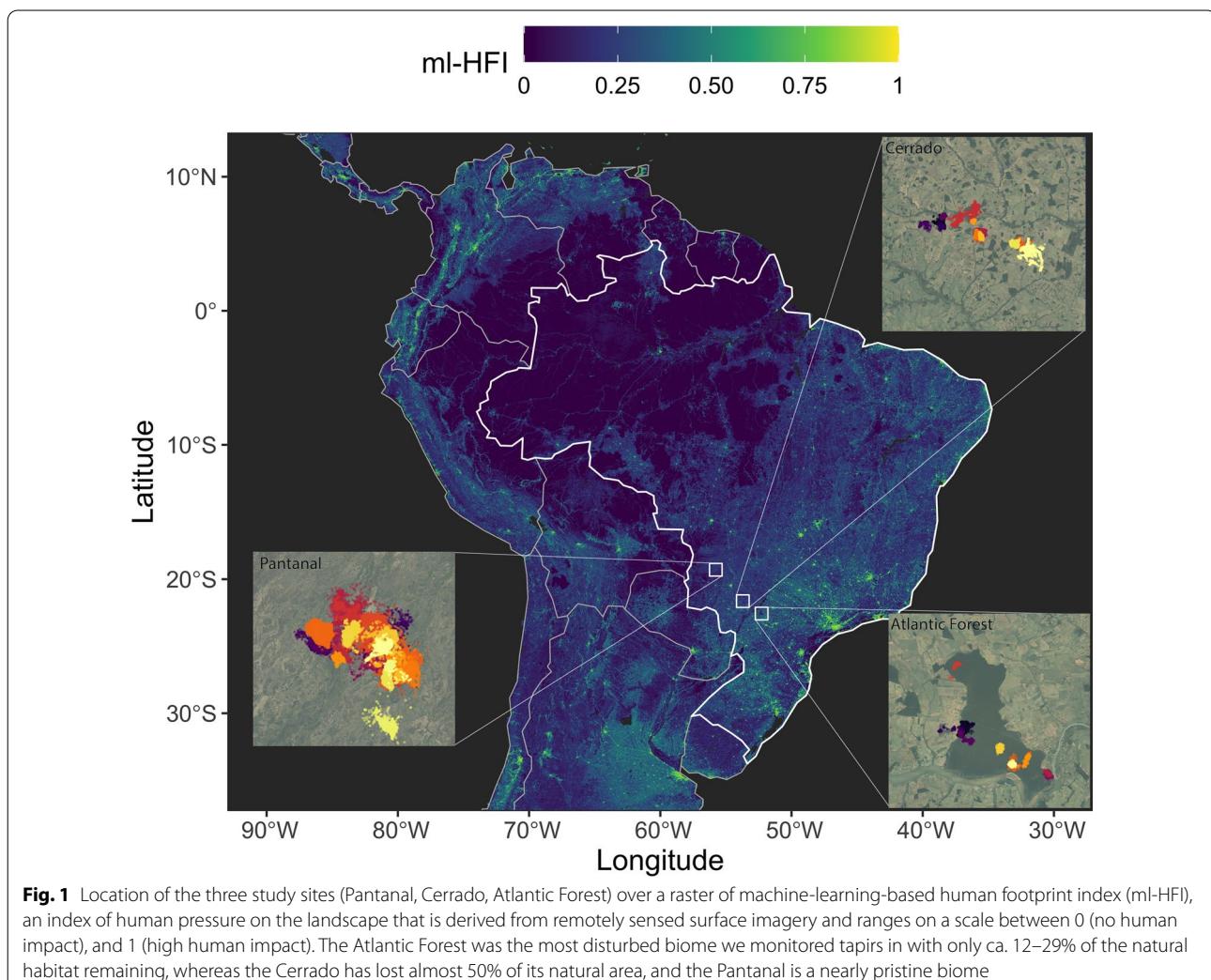
The data were collected in three different biomes in southern Brazil (Fig. 1): Atlantic Forest (1997–2007), Pantanal (2008–2019), and south-western Cerrado (2016–2018).

### Atlantic forest

Morro do Diabo State Park is a protected area located in the Municipality of Teodoro Sampaio (22°32'S, 52°18'W), state of São Paulo, in the southeastern region of Brazil. The park has an area of 370 km<sup>2</sup> composed of a mosaic of mature and secondary deciduous forest, surrounded by the Paranapanema River in the south, and of a matrix of cattle ranches and agriculture, mostly sugar cane, in the remaining borders [73]. Its average annual temperature is 22 °C and annual rainfall is 1347 mm [28]. The park is part of the “Planalto Forest,” which is distinguished from the coastal forest of the Atlantic Forest biome by having lower annual rainfall and a marked dry season from May to September and is thus more similar to the Cerrado biome [61]. In fact, the semi-deciduous forests of the “Planalto Forest” are similar to those occurring within or on the edges of the Cerrado [61]. The Atlantic Forest was the most disturbed biome we monitored tapirs in, with only ca. 12–29% of the natural habitat remaining [37, 58, 59, 66].

### Cerrado

The study site in the Cerrado biome is a 2200 km<sup>2</sup> mosaic of private properties (cattle ranches and farms) and landless people settlements within the Municipalities of Nova Alvorada do Sul and Nova Andradina, Mato Grosso do Sul State (21°37'S, 53°40'W). The area includes small fragments of natural Cerrado habitat (Cerradão fragments, gallery forests, and marshland—25% of the study area), surrounded by areas highly impacted by human activities such as agriculture (particularly sugarcane, soybean and corn), cattle-ranching (cultivated pastureland), eucalyptus plantations, rural communities, and highways. The average annual temperature is 25 °C and annual rainfall is



1185 mm. The Cerrado has lost almost 50% of its natural area due primarily to human driven land-use change to agriculture and cattle ranching [36, 45].

#### Pantanal

Baía das Pedras Ranch, a private property of 145 km<sup>2</sup>, is located in the Nhecolândia Sub-Region of the Southern Pantanal, Municipality of Aquidauana (19°20'S, 55°43'W), Mato Grosso do Sul State, in the central-western region of Brazil. The ranch includes a mosaic of seasonally inundated grasslands, lakes, gallery forests, scrub, and deciduous forests that supports an abundance of wildlife and is situated far away from the edges of the biome where deforestation and other anthropogenic threats are occurring. Cattle are raised extensively on the native grasses. The average annual temperature is 25 °C and annual rainfall is 1185 mm [12]. The Pantanal is a nearly pristine biome with substantially less human disturbance than the Atlantic Forest and Cerrado.

In each study site, tapirs were captured by darting after physical restraint in either box traps or pitfall traps, or by darting from a distance [55]. Animals were anesthetized mostly using a combination of butorphanol, medetomidine and ketamine, as described by Medici et al. [41] and Fernandes-Santos et al. [20]. Reversal agents were administrated at the end of procedures. The procedures carried out during immobilization included the subcutaneous insertion of a microchip, morphometric measurements, sex and age class determination, physical examination, collection of biological samples for health and genetic studies, and placement of a telemetry collar on adults. Animals were tracked using VHF tracking (all three regions, Telonics® MOD500) and GPS tracking (Pantanal and Cerrado; Telonics® TGW SOB and GPS IRIDIUM models). A total of 74 tapirs were tracked starting in July of 1997 until October of 2019, with the majority of the individuals being

in the Pantanal (46), while 17 and 11 were from the Cerrado and Atlantic Forest regions, respectively.

Tapirs equipped with VHF collars were monitored for 5 days per month with data collection concentrated during crepuscular times, 3 h at dawn (04:00–07:00 h) and 3 h at dusk (17:00–20:00 h). These periods are the two main peaks of tapir activity [40]. Each tapir was located every 30 min during the sampling periods. GPS collars were programmed to obtain a fix every hour and operated for a median of 15.4 months across all tagged tapirs. GPS fix success rates were 75% in the Pantanal and 90% in the Cerrado. The full dataset comprised 232,622 location estimates collected over a period of 22 years (for full details see Additional file 1: File S1). In addition to the tapir location data, we collected 883 and 174 measurements from tags in fixed locations in the Pantanal and Cerrado, respectively in order to calibrate the measurement error of the GPS tracking collars.

### Data analysis

Initial exploratory analyses were carried out in ctmmweb (version 0.2.11, [10]). All formal statistical analysis and plotting were performed using R (version 4.0.5, R Core Team 2021 [56]), with the packages ctmm (version 0.6.1, [9]), mgcv (version 1.8-36, [80]), ggplot2 (version 3.3.4, [79]), ggmap (version 3.0.0, [33]). The furrr package (version 0.2.2, [75]) was used for parallel computation on Windows machines. All R code can be found in the GitHub repository at <https://github.com/StefanoMezini/tapirs>. Details on the analyses are presented in Additional file 2: Appendix S2.

### Data calibration and cleaning

Before analysis, we performed an error calibration and data cleaning process to minimise the impacts of GPS measurement error and outliers on our subsequent analyses [21]. Data cleaning and calibration were carried out using the methods implemented in the ctmm R package. For this process, measurement error for location estimates collected via VHF telemetry was assumed to be insubstantial relative to the coarsely sampled movement data (median step length: 260.7 m) and raw locations were carried forward in the analyses. Measurement error on the GPS data was calibrated using a unitless Horizontal Dilution of Precision (HDOP), which quantifies the accuracy of each positional fix. We then estimated an equivalent range error with the HDOP values from the tags in fixed locations. This allowed for the unitless HDOP values to be converted into estimates of measurement error in meters. Notably, we found no effect of sampling technique, and thus error handling protocols, on home range area estimates (see Additional file 1: Appendix S1). After calibration, data points were

considered as outliers (and removed) if they had a large (error-informed) distance from the median location and the minimum speed required to explain the displacement was unusually high ( $\geq 1\text{m/s}$ ). The Atlantic Forest dataset contained a total of 4,082 observations, 8 (ca. 0.2%) of which were removed as outliers; the Pantanal dataset contained 139,138 observations, 914 (ca. 0.7%) of which were removed; while the Cerrado dataset contained 90,402 observations, 193 (ca. 0.2%) of which were removed.

### Movement modelling and home range estimation

For each of the monitored tapirs we quantified a number of key movement metrics and home range-related characteristics that allowed us to test for an effect of habitat composition and human disturbance on tapir movement behaviour. For this we first identified the best Continuous-Time Movement Model (CTMM) for each animal using the ctmm.select function from the ctmm package. This fits a series of CTMMs to location data using perturbative Hybrid Residual Maximum Likelihood (pHREML, [22]) and chooses the best model using small-sample-sized corrected Akaike's Information Criterion (AICc). The models used here are insensitive to sampling frequency (Johnson et al. 2008; [7, 24]) and they account for spatio-temporal autocorrelation in the data (when necessary), so they are robust to irregular or frequent sampling frequency [23]. The parameter estimates from each individual's movement model provided information on the tapir's home range crossing time ( $\tau_p$ , in days), and directional persistence timescale ( $\tau_v$ , in hours).

We then conditioned on the selected CTMMs to estimate each animal's 95% home range (HR) area (in  $\text{km}^2$ ) using small-sample-size bias corrected Autocorrelated Kernel Density Estimation (AKDE, [25, 49]), and average daily movement speed (in  $\text{km/day}$ ) using continuous-time speed and distance (CTSD) estimation [50].

### Movement pattern analyses

We were first interested in understanding whether home-range areas and movement metrics differed across the three biomes, as well as between animals of different age and sex. For these comparisons, home-range estimates were compared using the meta-analysis methods implemented in the ctmm package, which treats the individual home-range area estimates as having a chi-squared sampling distribution, and the population of home-range areas as having an inverse-Gaussian distribution [26]. This approach also allowed for uncertainty in the individual home-range estimates to be propagated to the population-level estimates. Home-range crossing time, directional persistence, and mean movement speed were analysed using generalized linear models (GLMs) with

a Gamma distribution and a log link function for the response. The GLMs were fit using the mgcv package [80] and Residual Maximum Likelihood (REML). Notably, because different tracking technologies were used to collect the tapir location data, we assessed whether tracking technique in-and-of-itself could have impacted the individual home range estimates. From these analyses found no effect of sampling technique on home range area (see Additional file 1: Appendix S1 for full details).

To test whether tapirs responded to different environment types, the HR sizes and average daily speeds were regressed against the proportions of the habitat types in each HR. For the Atlantic Forest, we used the habitat map provided in the park's management plan [28]. For the Pantanal and Cerrado, we obtained satellite imagery from the periods of data collection. Habitat classification was then carried out using GIS software, and a team of researchers confirmed the classifications in the field. The primary habitat types included: (1) forest, (2) savannah, (3) exposed soil, (4) floodplain, and (5) water. For full details on the habitat composition of the different study areas see Additional file 1: Figure S2 in Appendix S1. Similarly, the HR sizes and average daily speeds were regressed against their HR's average machine-learning-based human footprint index (ml-HFI) [34] to test whether human activity significantly altered the animals' behavior. Briefly, convolutional neural networks, are used to identify patterns of human activity from the Hansen Global Forest Change imagery version 1.7 (GFCv1.7, for full details see [34]). The resulting ml-HFI is an index of human pressure on the landscape that is derived from remotely sensed surface imagery and ranges on a scale between 0 (no human impact), and 1 (high human impact). For these models we applied Generalized Additive Models (GAMs) with a Gamma distribution and a log link function for the response. The Gamma distribution allows for more accurate significance testing and is an appropriate distribution for variables that range between 0 and  $\infty$ , while the log link scale allows HFI to have a multiplicative effect on the response. The GAMs were fit using the mgcv package [80] and REML, and the best model was selected using AICc based model selection. All analyses were carried out at both the 95% and 50% quantiles. The findings were consistent between quantiles and only results at the 95% quantile are presented in the main text. Full results for the 50% core home ranges are presented in Additional file 2: Appendix S2.

## Results

### Individual variation in movement and space use

The mean home-range size across all monitored tapirs was  $8.31 \text{ km}^2$  (95% CI: 6.53–10.42; Fig. 2), ranging between  $1.0 \text{ km}^2$  and  $29.7 \text{ km}^2$  (Fig. 3a). Tapirs had HR

crossing times of 0.72 days on average (95% CI 0.42–1.25), ranging from 0.05 to 12.8 days (Fig. 3b), and a mean velocity autocorrelation timescale of 0.44 h (95% CI 0.38–0.51), ranging from 0.17 to 1.88 h (Fig. 3c). We estimated that tapirs had mean movement speeds of 11.2 km/day (95% CI 10.1–12.3), ranging from 1.51 to 25.96 km/day (Fig. 3d). There was no evidence that average daily speed differed between sexes (females: 10.5 km/day, 95% CI 9.19–12.0; males: 11.9 km/day; 95% CI 10.3–13.7,  $p = 0.22$ , 4a), nor between age groups (adults: 11.8 km/day, 95% CI 10.6–13.2; sub-adults: 9.5 km/day, 95% CI 7.9–11.4;  $p = 0.053$ , Fig. 4b).

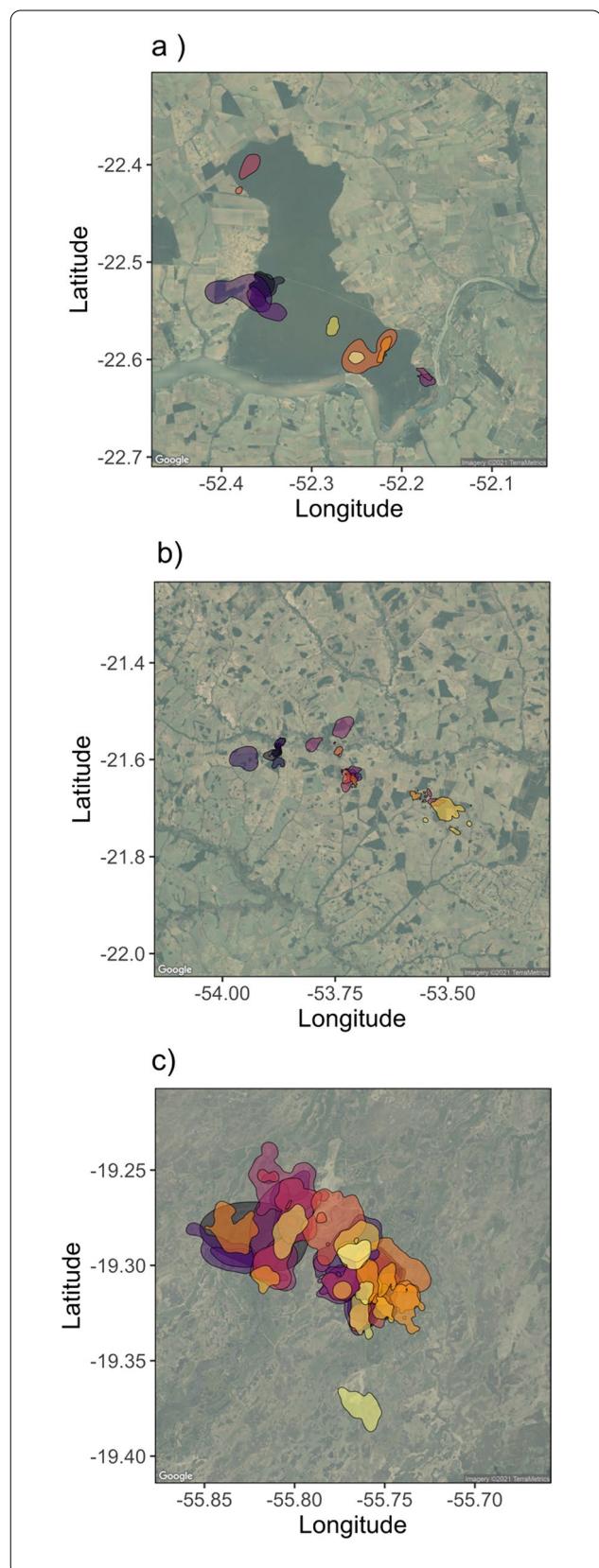
There was no evidence that home range sizes differed between sexes (males:  $5.46 \text{ km}^2$ , 95% CI 4.03–7.23; females:  $6.11 \text{ km}^2$ , 95% CI: 4.53–8.07 Fig. 4c), nor between age groups (adults:  $5.37 \text{ km}^2$ , 95% CI: 4.39–11.64; sub-adults:  $6.98 \text{ km}^2$ , 95% CI: 3.87–11.65; Fig. 4d). We estimated the male/female ratio of mean home-range areas to be 0.87 (0.56–1.30), and the adult/subadult ratio to be 0.70 (0.37–1.32), both of which include 1 and are thus non-significant.

### Variation in movement across biomes and variation in human disturbance

The Atlantic Forest, Cerrado, and Pantanal vary substantially in habitat composition, levels of human disturbance, and tapir population densities. Despite these differences, we found that lowland tapir movement behaviour and space use were consistent across all three biomes (Fig. 3).

We also found that habitat type had little effect on HR area or average individual movement speeds. The best HR area regression model only accounted for the effect of areas of exposed soil (approximate p-value: 0.023,  $R^2_{adj} = 0.48$ ; Fig. 5a), while no land use types had a significant effect on an animal's average speed. There was very little difference between the AIC of the full model (315.69,  $df=10.18$ , 7 predictors and an intercept) and that of the intercept-only model (310.89,  $df=2$ ). However, the directional persistence term ( $\tau_v$ ) was marginally, though non-significantly lower for animals who had a higher amount of forested area ( $p=0.093$ ; Fig. 5b) or water ( $p=0.025$ ) in their home ranges. Importantly, we note here that the significant differences in directional persistence persisted even after adjusted for the increased location error in the forested areas.

HFI had no significant effect on lowland tapir home range size ( $p=0.90$ ; Fig. 6a), nor average daily movement speed ( $p=0.53$ ; Fig. 6b), nor directional persistence ( $p=0.596$ ,  $R^2_{adj} = -0.0184$ ). A tapir living in a near pristine environment (HFI=0.004) had a home range estimate of  $7.77 \text{ km}^2$  (95% CI 2.12–28.6) and an average speed of 13.2 km/day (95% CI 7.8–22.1) with a



**Fig. 2** Autocorrelated kernel density estimates of each tapir's 95% home range in each of the three regions: **a** Atlantic forest, **b** Cerrado, and **c** pantanal

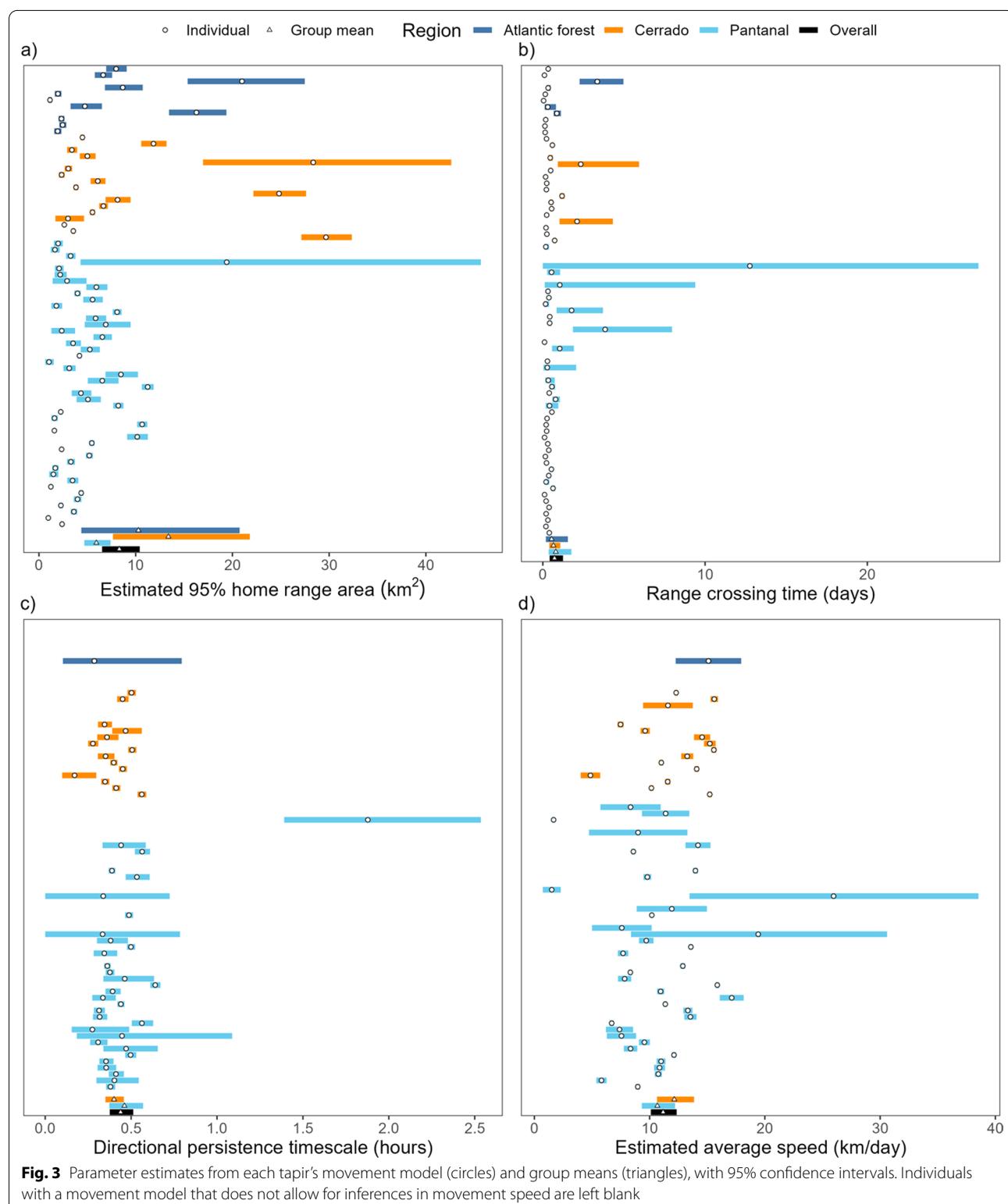
directional persistence of 0.36 h (95% CI 0.16–0.78), while a tapir from the most altered habitat we monitored ( $HFI=0.31$ ) had an estimated home range area of  $6.93 \text{ km}^2$  (95% CI 3.36–14.30) and an average speed of 10.4 km/day (95% CI 8.3–13.2) with a directional persistence of 0.48 h (95% CI 0.34–0.68).

## Discussion

Understanding individual movement and space use requirements is a key step in conservation planning [3]. Prior to the present study, very little was known about the movement ecology of tapirs despite their vulnerable status and declining population sizes [74]. From detailed tracking of 74 tapirs collected over 22 years, we found that although individuals varied in their movement, these inter-individual differences were not explained by differences in age, sex, habitat composition, biome, nor human disturbance. Overall, human activity and land use change did not appear to significantly affect their behaviour one way or another. This contradicts patterns in large herbivores generally [16, 71], and further emphasizes the need to understand the movement ecology of target populations when designing conservation and recovery strategies.

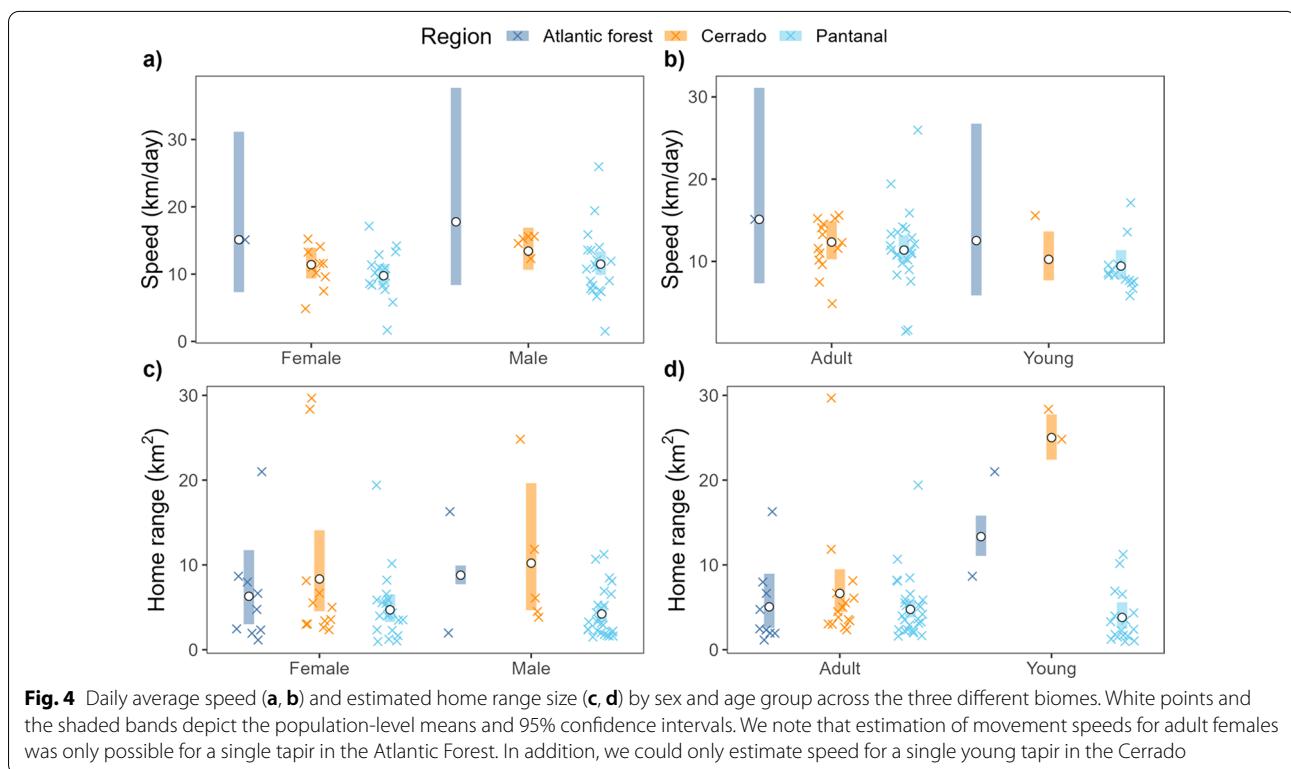
### The ecology of lowland tapir space use

Interestingly, we found that the home range sizes and mean daily movement speeds of lowland tapirs were consistent across the three study sites. This consistency in movement was surprising as these different biomes have substantially different habitat compositions, patterns of seasonality, and productivity [47], see also Additional file 1: Appendix S1). Tapirs living in the Pantanal, for instance, occupy a near pristine ecosystem but must cope with significant seasonal flooding, whereas individuals in the Cerrado occupied an agricultural and cattle ranching mosaic with more stability across seasons. The unique requirements of these three different biomes, however, did not impact the space use and movement speed of tapirs in any statistically detectable way. Furthermore, the only pre-existing study on tapir movement found that individuals had complex home range structures, with multiple core areas of use that were established according to the distribution of patches of preferred habitat types [69]. While individuals may exhibit differential use of patchily distributed resources, we found that habitat composition had no effect on home range sizes. In addition to exhibiting little inter-individual variation in



movement, variogram analysis [24] showed that tapir movement was extremely consistent over time (see also [22]. Here again, this seasonal stability in movement

was interesting, especially for animals living in the Pantanal where, every year, large parts of the biome change from terrestrial into aquatic habitats and vice-versa [1].



We note though that the flooding regime in the Pantanal has been changing over the last decade and the biome is expected to become drier under the IPCC's climate change scenarios [38].

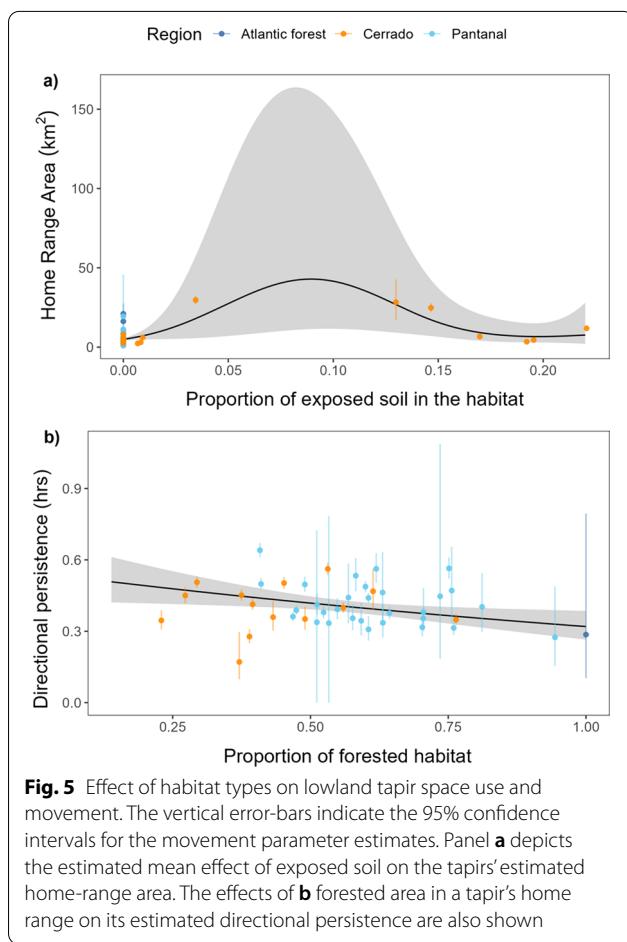
We did find that animals with a higher proportion of forest and/or more water bodies in their home ranges had reduced directional persistence. This shows how habitat complexity can impact movement [15], with potential implications for foraging efficiency and encounter rates [5, 39, 77]. Nonetheless, these differences did not translate into patterns in tapir home range sizes and mean daily movement speeds.

#### Lowland tapir movement and human disturbance

This is the first study aimed at understanding how lowland tapir space use and movement vary across differing biomes and degrees of human disturbance. Contrary to our initial expectations, and to patterns in large herbivores generally [16], human impacts on the landscape had no measurable effect on tapir movement. To put this landscape scale effect into perspective, tapirs inhabiting the Atlantic Forest, the most disturbed biome with only ca. 12–29% of habitat remaining [37, 58, 59, 66], had home ranges that were comparable in size to tapirs inhabiting the Cerrado, a biome that has lost almost 50% of its natural area (36, 45), and the Pantanal, a near pristine biome. Notably, the Lowland Tapir Conservation Action Plan

published by the IUCN SSC Tapir Specialist Group (TSG) in 2007 [43], and the Lowland Tapir National Action Plan (PAN—Plano de Ação Nacional, ICMBIO—Instituto Chico Mendes de Conservação da Biodiversidade, Brazil) published in 2019 prioritize the mitigation of the impacts of small, isolated tapir populations. Population isolation thus emerges as one of the most important threats to the species' long-term persistence. However, addressing this issue will require additional efforts as the average and maximum distances we recorded for tapir movements were substantially less than the distances between most tapir populations.

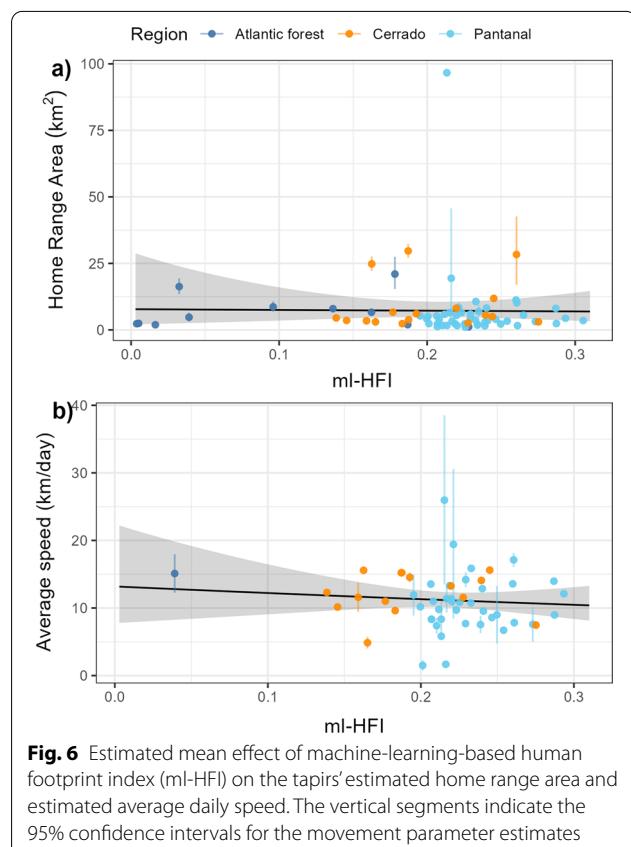
Humans are directly responsible for more than one-quarter of global terrestrial vertebrate mortality [30]. Mortality at this scale is expected to impose strong selection pressure on animal populations [53, 67]. As genotypic adaptation takes generations to occur [4], behavioral plasticity provides the most immediate response to the pressures of Human Induced Rapid Environmental Change (HIREC, [65]). The capacity for behavioural plasticity in movement and space use in response to human disturbance is especially important for long-lived, K-selected species such as tapirs [46, 60, 65] that take years to reach sexual maturity and have long inter-generational intervals [40]. Despite the key importance of behavioural adaptations in response to HIREC, tapir movement appeared to exhibit very little plasticity in



response to human disturbance. The lack of any measurable response to human activity suggests that tapirs living near humans may experience increased exposure to vehicle collisions [2, 42], pesticide and environmental pollutants [20, 41, 44] and poaching [62]. Human modified habitats thus risk being ecological traps [64] for tapirs as individuals showed no detectable responses to degradations in habitat quality. Although tapir home range area and mean daily movement speed exhibited no statistically detectable response to the human footprint index, it is possible that individuals are responding to human disturbance at a finer temporal and/or spatial scale than the long-term averages that were examined here. It may also be possible that tapirs exhibit non-linear, or even binary, responses to human disturbance that were not possible to detect. Future investigation into lowland tapir behaviour in more heavily modified habitats is clearly warranted.

## Conclusions

We compared home range areas and movement behavior of lowland tapirs using telemetry data collected over 22 years across 3 biomes in southern Brazil: the



**Fig. 6** Estimated mean effect of machine-learning-based human footprint index (ml-HFI) on the tapirs' estimated home range area and estimated average daily speed. The vertical segments indicate the 95% confidence intervals for the movement parameter estimates

Pantanal, Cerrado, and Atlantic Forest. These data represent the largest lowland tapir tracking dataset yet to be collected, with over 232,000 locations from 74 tracked individuals and fill a critical knowledge gap in lowland tapir ecology, which can contribute to long-term species management and conservation planning. Contrary to our expectations, we observed very little individual variability in lowland tapir space use and movement, and human impacts on the landscape also had no measurable effect on their movement. Lowland tapir movement behaviour thus appears to exhibit very little phenotypic plasticity. The lack of any adaptive response to anthropogenic disturbance suggests that human modified habitats risk being ecological traps for tapirs and this information should be factored into conservation actions aimed towards protecting lowland tapir populations.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40462-022-00313-w>.

**Additional file 1:** Details on the habitat composition of the different biomes.

**Additional file 2:** Details on the R scripts used to generate the results presented in the main text.

## Acknowledgements

The study of tapir movement ecology has been an important component of the long-term activities of the Lowland Tapir Conservation Initiative (LTCI) – Instituto de Pesquisas Ecológicas (IPÊ) in Brazil. The LTCI has the institutional support from the International Union for Conservation of Nature (IUCN) Species Survival Commission (SSC) Tapir Specialist Group (TSG), Association of Zoos and Aquariums (AZA) Tapir Taxon Advisory Group (TAG), and European Association of Zoos and Aquariums (EAZA) Tapir Taxon Advisory Group (TAG). LTCI's financial support comes from national and international agencies, including zoological institutions, foundations, private businesses, and private individuals. EPM would like to thank the Smithsonian Conservation Biology Institute (SCBI) for hosting her for a 2-month research visit for initial data processing and analysis.

## Authors' contributions

EPM devised the field study, collected the data, and performed preliminary analyses. EPM, JMC, and MJN conceived the scope of the project and analytical pipeline. SM, CHF, and MJN contributed to data analysis. All authors contributed to the writing. All authors read and approved the final manuscript.

## Funding

MJN was supported by an NSERC Discovery Grant RGPIN-2021-02758. This work was partially funded by the Center of Advanced Systems Understanding (CASUS) which is financed by Germany's Federal Ministry of Education and Research (BMBF) and by the Saxon Ministry for Science, Culture and Tourism (SMWK) with tax funds on the basis of the budget approved by the Saxon State Parliament. CHF and JMC were supported by NSF IIBR 1915347.

## Availability of data and materials

All R code and data necessary to reproduce the analyses in the main text can be found in the GitHub repository at <https://github.com/StefanoMezzini/tapirs>.

## Declarations

### Ethics approval and consent to participate

The Instituto Chico Mendes de Conservação da Biodiversidade (ICMBIO) provided the required annual permits for the capture and immobilization of tapirs and collection of biological samples (SISBIO# 14,603). The Comissão Técnico-Científica (COTEC) do Instituto Florestal do Estado de São Paulo (IF-SP) provided the required permit to carry out research in Morro do Diabo State Park (SMA# 40624/1996). All protocols for the capture, anesthesia, handling and sampling of tapirs have been reviewed and approved by the Veterinary Advisors of the Association of Zoos and Aquariums (AZA) – Tapir Taxon Advisory Group (TAG), and the Veterinary Committee of the IUCN SSC Tapir Specialist Group (TSG).

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Lowland Tapir Conservation Initiative (LTCI), Instituto de Pesquisas Ecológicas (IPÊ), Rodovia Dom Pedro I, km 47, Nazaré Paulista, São Paulo 12960-000, Brazil.

<sup>2</sup>IUCN SSC Tapir Specialist Group (TSG), Campo Grande, Brazil. <sup>3</sup>Escola Superior de Conservação Ambiental E Sustentabilidade (ESCAS/IPÊ), Rodovia Dom Pedro I, km 47, Nazaré Paulista, São Paulo 12960-000, Brazil. <sup>4</sup>The Irving K. Barber Faculty of Science, The University of British Columbia, Okanagan Campus, Kelowna, Canada. <sup>5</sup>University of Maryland College Park, College Park, MD, USA.

<sup>6</sup>Smithsonian Conservation Biology Institute, Front Royal, VA, USA. <sup>7</sup>Center for Advanced Systems Understanding (CASUS), Görlitz, Germany. <sup>8</sup>Helmholtz-Zentrum Dresden Rossendorf (HZDR), Dresden, Germany. <sup>9</sup>Department

of Ecological Modelling, Helmholtz Centre for Environmental Research (UFZ), Leipzig, Germany.

Received: 17 November 2021 Accepted: 2 March 2022

Published online: 14 March 2022

## References

- Alho CJR. Biodiversity of the pantanal: response to seasonal flooding regime and to environmental degradation. *Braz J Biol.* 2008. <https://doi.org/10.1590/S1519-69842008000500005>.
- Abra FD, da Costa CA, Garbino GST, Medici EP. Use of unfenced highway underpasses by lowland tapirs and other medium and large mammals in central-western Brazil. *Perspect Ecol Conserv.* 2020;18:247–56. <https://doi.org/10.1016/j.pecon.2020.10.006>.
- Allen AM, Singh NJ. Linking movement ecology with wildlife management and conservation. *Front Ecol Evol.* 2016. <https://doi.org/10.3389/fevo.2015.00155>.
- Barnosky AD, Kraatz BP. The role of climatic change in the evolution of mammals. *Bioscience.* 2007. <https://doi.org/10.1641/B570615>.
- Bartumeus F, Catalan J, Viswanathan GM, Raposo EP, Da Luz MGE. The influence of turning angles on the success of non-oriented animal searches. *J Theor Biol.* 2008. <https://doi.org/10.1016/j.jtbi.2008.01.009>.
- Bauer S, Hoye BJ. Migratory animals couple biodiversity and ecosystem functioning worldwide. *Science.* 2014. <https://doi.org/10.1126/science.124552>.
- Blackwell PG, Niu M, Lambert MS, LaPoint SD. Exact Bayesian inference for animal movement in continuous time. *Methods Ecol Evol.* 2016. <https://doi.org/10.1111/2041-210X.12460>.
- Brooks TM, Mittermeier RA, Mittermeier CG, Fonseca GAB, Rylands AB, Konstant WR, et al. Habitat loss and extinction in the hotspots of biodiversity. *Conserv Biol.* 2002. <https://doi.org/10.1046/j.1523-1739.2002.00530.x>.
- Calabrese JM, Fleming CH, Gurarie E. Ctmr: an r Package for analyzing animal relocation data as a continuous-time stochastic process. *Methods Ecol Evol.* 2016. <https://doi.org/10.1111/2041-210X.12559>.
- Calabrese JM, Fleming CH, Noonan MJ, Dong X. Ctmmweb: a graphical user interface for autocorrelation-informed home range estimation. *Wildl Soc Bull.* 2021. <https://doi.org/10.1002/wsb.1154>.
- Calder WA III. Ecological scaling: mammals and birds. *Annu Rev Ecol Syst.* 1983;14:213–30.
- Calheiros DF, Fonseca Junior WC. Perspectivas de estudos ecológicos sobre o Pantanal. Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA-CPAP). Corumbá, Mato Grosso do Sul, Brazil. 1996.
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, et al. Biodiversity loss and its impact on humanity. *Nature.* 2012. <https://doi.org/10.1038/nature11148>.
- Cosgrove AJ, McWhorter TJ, Maron M. Consequences of impediments to animal movements at different scales: a conceptual framework and review. *Divers Distrib.* 2018. <https://doi.org/10.1111/ddi.12699>.
- Dickie M, Serrouya R, McNay RS, Boutin S. Faster and farther: wolf movement on linear features and implications for hunting behaviour. *J Appl Ecol.* 2017. <https://doi.org/10.1111/1365-2664.12732>.
- Doherty TS, Hays GC, Driscoll DA. Human disturbance causes widespread disruption of animal movement. *Nat Ecol Evol.* 2021. <https://doi.org/10.1038/s41559-020-01380-1>.
- Esfahani HS, Ramirez MT. Institutions, infrastructure, and economic growth. *J Dev Econ.* 2003. [https://doi.org/10.1016/S0304-3878\(02\)00105-0](https://doi.org/10.1016/S0304-3878(02)00105-0).
- Fahrig L. Relative effects of habitat loss and fragmentation on population extinction. *J Wildl Manag.* 1997;61(3):603–10.
- Fahrig L. Non-optimal animal movement in human-altered landscapes. *Funct Ecol.* 2007. <https://doi.org/10.1111/j.1365-2435.2007.01326.x>.
- Fernandes-Santos RC, Medici EP, Testa-Jose C, Micheletti T. Health assessment of wild lowland tapirs (*Tapirus terrestris*) in the highly threatened Cerrado biome. *Braz J Wildl Dis.* 2020. <https://doi.org/10.7589/2018-10-244>.
- Fleming CH, Drescher-Lehman J, Noonan MJ, Akre TSB, Brown DJ, Cochrane MM, et al. A comprehensive framework for handling location error in animal tracking data. *Ecology.* 2020.

22. Fleming CH, Noonan MJ, Medici EP, Calabrese JM. Overcoming the challenge of small effective sample sizes in home-range estimation. *Methods Ecol Evol.* 2019. <https://doi.org/10.1111/2104-210X.13270>.
23. Fleming CH, Sheldon D, Fagan WF, Leimgruber P, Mueller T, Nandintsetseg D, et al. Correcting for missing and irregular data in home-range estimation. *Ecol Appl.* 2018. <https://doi.org/10.1002/eam.1704>.
24. Fleming CH, Calabrese JM, Mueller T, Olson KA, Leimgruber P, Fagan WF. From fine-scale foraging to home ranges: a semivariance approach to identifying movement modes across spatiotemporal scales. *Am Nat.* 2014. <https://doi.org/10.1086/675504>.
25. Fleming CH, Calabrese JM. A new kernel density estimator for accurate home-range and species-range area estimation. *Methods Ecol Evol.* 2017. <https://doi.org/10.1111/2104-210X.12673>.
26. Fleming CH, Deznabi I, Alavi S, Crofoot MC, Hirsch BT, Medici EP, et al. Population-level inference for home-range areas. *bioRxiv.* 2021. <https://doi.org/10.1101/2021.07.05.451204>.
27. Fundação SOS Mata Atlântica. Atlas dos remanescentes florestais da Mata Atlântica: período 2000–2005. 2008. <http://mapas.sosma.org.br>. Accessed 05 Nov 2021.
28. Faria H, Pires AS. Parque Estadual Morro Do Diabo—Plano de Manejo. Governo do Estado de São Paulo, Secretaria do Meio Ambiente, Instituto Florestal. Santa Cruz do Rio Pardo, São Paulo, Brazil. 2006.
29. Gardner AL. Mammals of South America, vol. 1: Marsupials, Xenarthrans, Shrews, and Bats. University of Chicago Press; 2008.
30. Hill JE, DeVault TL, Belant JL. Cause-specific mortality of the world's terrestrial vertebrates. *Glob Ecol Biogeogr.* 2019. <https://doi.org/10.1111/geb.12881>.
31. Hill JE, DeVault TL, Wang G, Belant JL. Anthropogenic mortality in mammals increases with the human footprint. *Front Ecol Environ.* 2020. <https://doi.org/10.1002/fee.2127>.
32. Hooper DU, Adair EC, Cardinale BJ, Byrnes JEK, Hungate BA, Matulich KL, et al. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature.* 2012. <https://doi.org/10.1038/nature11118>.
33. Kahle D, Wickham H. Ggmap: spatial visualization with Ggplot2. *R J.* 2013;5(1):144.
34. Keys PW, Barnes EA, Carter NH. A machine-learning approach to human footprint index estimation with applications to sustainable development. *Environ Res Lett.* 2021. <https://doi.org/10.1088/1748-9326/abe00a>.
35. Lucherini M, Lovari S. Habitat richness affects home range size in the red fox *Vulpes vulpes*. *Behav Proc.* 1996. [https://doi.org/10.1016/0376-6357\(95\)00018-6](https://doi.org/10.1016/0376-6357(95)00018-6).
36. Machado RB, Ramos-Neto MB, Pereira PGP, Caldas EF, Goncalves DA, Santos NS, et al. Estimativas de perda da área do cerrado brasileiro. Relatório Técnico. Conservação Internacional, Brasília, DF, Brazil. 2004.
37. Mapbiomas 2021. Projeto MapBiomas - Coleção 6 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. <https://mapbiomas.org/download>.
38. Marengo JA, Oliveira GS, Alves LM. Climate change scenarios in the Pantanal. In: Bergier I, Assine ML, editors. *Dynamics of the Pantanal Wetland in South America*. Springer; 2016. p. 227–38.
39. Martinez-Garcia R, Fleming CH, Seppelt R, Fagan WF, Calabrese JM. How range residency and long-range perception change encounter rates. *J Theor Biol.* 2020. <https://doi.org/10.1016/j.jtbi.2020.110267>.
40. Medici EP. Family Tapiridae (Tapirs). In: Wilson DE, Mittermeier RA, editors. *Handbook of the mammals of the world*, vol. 2. Hoofed Mammals. Lynx Edicions Barcelona; 2011. p. 182–204.
41. Medici EP, Mangini PR, Fernandes-Santos RC. Health assessment of wild lowland tapir (*Tapirus terrestris*) populations in the Atlantic Forest and Pantanal Biomes, Brazil (1996–2012). *J Wildl Dis.* 2014. <https://doi.org/10.7589/2014-02-029>.
42. Medici EP, Abra FD. Licoes aprendidas na conservação da anta brasileira e os desafios para mitigar uma de suas ameaças mais graves: o atropelamento em rodovias. *Boletim da Sociedade Brasileira de Mastozoologia.* 2019;85:152–60.
43. Medici EP, Desbiez ALJ, Gonçalves da Silva A, Jerusalinsky L, Chassot O, Montenegro OL, et al. Lowland Tapir (*Tapirus terrestris*) conservation workshop: final report. IUCN SSC Tapir Specialist Group & IUCN SSC Conservation Planning Specialist Group (Brazil Network). 2007.
44. Medici EP, Fernandes-Santos RC, Testa-Jose C, Godinho AF, Brand AF. Lowland tapir exposure to pesticides and metals in the Brazilian Cerrado. *Wildl Res.* 2021. <https://doi.org/10.1071/WR19183>.
45. Ministério da Agricultura, Pecuária e Abastecimento. TerraClass – Gestão Integrada da Paisagem no Bioma Cerrado. 2021.
46. Montgomery RA, Macdonald DW, Hayward MW. The inducible defences of large mammals to human lethality. *Funct Ecol.* 2020. <https://doi.org/10.1111/1365-2435.13685>.
47. Morato RG, Stabach JA, Fleming CH, Calabrese JM, Paula RC, Ferraz KMPM, et al. Space use and movement of a neotropical top predator: the endangered jaguar. *PLoS ONE.* 2016. <https://doi.org/10.1371/journal.pone.0168176>.
48. Nilsen EB, Herfindal I, Linnell JDC. Can intra-specific variation in carnivore home-range size be explained using remote-sensing estimates of environmental productivity? *Ecoscience.* 2005. <https://doi.org/10.2980/i1195-6860-12-1-68.1>.
49. Noonan MJ, Tucker MA, Fleming CH, Alberts SC, Ali AH, Altmann J, et al. A comprehensive analysis of autocorrelation and bias in home range estimation. *Ecol Monogr.* 2019. <https://doi.org/10.1002/ecm.1344>.
50. Noonan MJ, Fleming CH, Akre TS, Drescher-Lehman J, Gurarie E, Harrison AL, et al. Scale-insensitive estimation of speed and distance traveled from animal tracking data. *Mov Ecol.* 2019. <https://doi.org/10.1186/s40462-019-0177-1>.
51. Noonan MJ, Fleming CH, Tucker MA, Kays R, Harrison AL, Crofoot MC, et al. Effects of body size on estimation of mammalian area requirements. *Conserv Biol.* 2020. <https://doi.org/10.1111/cobi.13495>.
52. Noss AJ, Cuellar RL, Barrientos J, Maffei L, Cuellar E, Arispe R, et al. A camera trapping and radio telemetry study of lowland tapir (*Tapirus terrestris*) in Bolivian dry forests. *Plant Divers.* 2003;229:44–5.
53. Oro D, Genovart M, Tavecchia G, Fowler MS, Martinez-Abrain A. Ecological and evolutionary implications of food subsidies from humans. *Ecol Lett.* 2013. <https://doi.org/10.1111/ele.12187>.
54. Powers RP, Jetz W. Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. *Nat Clim Chang.* 2019. <https://doi.org/10.1038/s41558-019-0406-z>.
55. Quse VB, Fernandes-Santos RC. Tapir Veterinary Manual. IUCN SSC Tapir Specialist Group (TSG). 2014.
56. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>. 2021.
57. Relyea RA, Lawrence RK, Demarais S. Home range of desert mule deer: testing the body-size and habitat-productivity hypotheses. *J Wildl Manag.* 2000;64(1):146–53.
58. Rezende CL, Scarano FR, Assad ED, Joly CA, Metzger JP, Strassburg BBN, Tabarelli M, Fonseca GA, Mittermeier RA. From hotspot to hotspot: An opportunity for the Brazilian Atlantic Forest. *Perspect Ecol Conserv.* 2018;16:208–14. <https://doi.org/10.1016/j.pecon.2018.10.002>.
59. Ribeiro MC, Metzger JP, Martensen AC, Ponsoni FJ, Hirota MM. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications. *Conserv Biol.* 2009;14:1141–53. <https://doi.org/10.1016/j.biocon.2009.02.021>.
60. Rosenheim JA, Tabashnik BE. Influence of generation time on the rate of response to selection. *Am Nat.* 1991;137(4):527–41.
61. Salis SM, Shepherd GJ, Joly CA. Floristic comparison of mesophytic semideciduous forests of the interior of the state of São Paulo, Southeast Brazil. *Vegetatio.* 1995;119(2):155–64.
62. Sanches A, Perez WAM, Figueiredo MG, Rossini BC, Cervini M, Galetti PM, et al. Wildlife forensic DNA and lowland tapir (*Tapirus terrestris*) poaching. *Conserv Genet Resour.* 2011. <https://doi.org/10.1007/s12686-010-9318-y>.
63. Schick RS, Loarie SR, Colchero F, Best BD, Boustany A, Conde DA, et al. Understanding movement data and movement processes: current and emerging directions. *Ecol Lett.* 2008. <https://doi.org/10.1111/j.1461-0248.2008.01249.x>.
64. Schlaepfer MA, Runge MC, Sherman PW. Ecological and evolutionary traps. *Trends Ecol Evol.* 2002. [https://doi.org/10.1016/S0169-5347\(02\)02580-6](https://doi.org/10.1016/S0169-5347(02)02580-6).
65. Sih A, Ferrari MCO, Harris DJ. Evolution and behavioural responses to human-induced rapid environmental change. *Evol Appl.* 2011. <https://doi.org/10.1111/j.1752-4571.2010.00166.x>.

67. Swaddle JP, Francis CD, Barber JR, Cooper CB, Kyba CCM, Dominoni DM, et al. A framework to assess evolutionary responses to anthropogenic light and sound. *Trends Ecol Evol.* 2015. <https://doi.org/10.1016/j.tree.2015.06.009>.
68. Thatte P, Chandramouli A, Tyagi A, Patel K, Baro P, Chhattani H, et al. Human footprint differentially impacts genetic connectivity of four wide-ranging mammals in a fragmented landscape. *Divers Distrib.* 2020. <https://doi.org/10.1111/ddi.13022>.
69. Tobler MW. The ecology of the lowland tapir in Madre de Dios, Peru: using new technologies to study large rainforest mammals. Texas A&M University; 2008.
70. Toews M, Juanes F, Burton AC. Mammal responses to the human footprint vary across species and stressors. *J Environ Manage.* 2018. <https://doi.org/10.1016/j.jenvman.2018.04.009>.
71. Tucker MA, Bohning-Gaese K, Fagan WF, Fryxell JM, Moorter BV, et al. Moving in the Anthropocene: global reductions in terrestrial mammalian movements. *Science.* 2018. <https://doi.org/10.1126/science.aam9712>.
72. Tucker MA, Busana M, Huijbregts MAJ, Ford AT. Human-induced reduction in mammalian movements impacts seed dispersal in the tropics. *Ecography.* 2021. <https://doi.org/10.1111/ecog.05210>.
73. Uezu A, Beyer DD, Metzger JP. Can agroforest woodlots work as stepping-stones for birds in the Atlantic Forest region? *Biodivers Conserv.* 2008. <https://doi.org/10.1007/s10531-008-9329-0>.
74. Varela D, Flesher K, Cartes JL, Bustos S, Chalukian S, Ayala G, et al. Tapirus terrestris. The IUCN Red List of Threatened Species. 2019. <https://www.iucnredlist.org/species/21474/45174127>. Accessed 05 Nov 2021.
75. Vaughan D, Dancho M. Furrr: apply mapping functions in parallel using futures. 2021.
76. Venter O, Brodeur NN, Nemiroff L, Belland B, Dolinsek IJ, Grant JWA. Threats to endangered species in Canada. *Bioscience.* 2006;56(11):903–10.
77. Visser AW, Kiorboe T. Plankton motility patterns and encounter rates. *Oecologia.* 2006. <https://doi.org/10.1007/s00442-006-0385-4>.
78. Wall J, Wittemyer G, Klinkenberg B, LeMay V, Blake S, Strindberg S, et al. Human footprint and protected areas shape elephant range across Africa. *Curr Biol.* 2021. <https://doi.org/10.1016/j.cub.2021.03.042>.
79. Wickham H. *Ggplot2: elegant graphics for data analysis.* New York: Springer; 2016.
80. Wood SN. *Generalized additive models: an introduction with R.* 2nd edn. Chapman; Hall/CRC; 2017.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)





a multidisciplinary open access journal  
*published by Canadian Science Publishing*

**Assessing the representation of species included within the  
 Canadian Living Planet Index**

Journal:	<i>FACETS</i>
Manuscript ID:	facets-2022-0063
Manuscript Type:	Research Article
Date Submitted by the Author:	17-Mar-2022
Complete List of Authors:	Currie, Jessica; World Wildlife Fund, Science, Research & Innovation; Burant, Joseph; Université de Montréal, Département de sciences biologiques; McGill University, Department of Biology; Canadian Institute of Ecology and Evolution, Living Data Project Marconi, Valentina; Zoological Society of London, Indicators and Assessments Unit; Imperial College London, Department of Life Sciences Blain, Stephanie; The University of British Columbia, Department of Zoology and Biodiversity Research Centre Emry, Sandra; The University of British Columbia, Department of Zoology and Biodiversity Research Centre Hebert, Katherine; Université de Sherbrooke, Département de biologie Xie, Garland; University of Toronto, Department of Biological Sciences Moore, Nikki; McGill University, Department of Biology Wang, Xueqi; University of Guelph, Department of Integrative Biology Brown, Andrea; McGill University, Department of Biology Grevstad, Lara; The University of British Columbia, Department of Geography McRae, Louise; Institute of Zoology of the Zoological Society of London, Indicators and Assessments Unit Mezzini, Stefano; The University of British Columbia, Department of Biology Pata, Patrick; The University of British Columbia, Department of Earth, Ocean, and Atmospheric Sciences Freeman, Robin; Institute of Zoology of the Zoological Society of London, Indicators and Assessments Unit
Primary manuscript subject:	Conservation and Sustainability < Integrative Sciences
Secondary manuscript subject:	Ecology and Evolution < Biological and Life Sciences
Is the paper being submitted for consideration in a FACETS Collection? :	Not Applicable (regular submission)

SCHOLARONE™  
Manuscripts

1   **Title:** Assessing the representation of species included within the Canadian Living Planet Index

2

3   **Authors:** Jessica Currie<sup>a</sup>, Joseph B. Burant<sup>b,c,d</sup>, Valentina Marconi<sup>e,f</sup>, Stephanie A. Blain<sup>g</sup>, Sandra Emry<sup>g</sup>,  
4   Katherine Hébert<sup>h</sup>, Garland Xie<sup>i</sup>, Nikki Moore<sup>b</sup>, Xueqi Wang<sup>j</sup>, Andrea Brown<sup>b</sup>, Lara Grevstad<sup>k</sup>, Louise  
5   McRae<sup>e</sup>, Stefano Mezzini<sup>l</sup>, Patrick Pata<sup>m</sup>; Robin Freeman<sup>j</sup>

6

7   **Affiliation:**

8   <sup>a</sup> World Wildlife Fund Canada, 410 Adelaide Street West, Toronto ON M5V 1S8; <sup>b</sup> Department of Biology,  
9   McGill University, 1205 Docteur Penfield Avenue, Montreal QC H3A 1B1, Canada; <sup>c</sup> Département de  
10   sciences biologiques, Université de Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal QC H2V 0B3,  
11   Canada; <sup>d</sup> Living Data Project, Canadian Institute of Ecology and Evolution, Vancouver BC, Canada; <sup>e</sup>  
12   Indicators and Assessments Unit, Institute of Zoology, Zoological Society of London, Regent's Park,  
13   London NW1 4RY, United Kingdom; <sup>f</sup> Department of Life Sciences (Silwood Park), Imperial College  
14   London, Buckhurst Road, Ascot, Berkshire SL5 7PY, United Kingdom; <sup>g</sup> Department of Zoology and  
15   Biodiversity Research Centre, University of British Columbia, 6270 University Boulevard, Vancouver BC  
16   V6T 1Z4, Canada; <sup>h</sup> Département de biologie, Université de Sherbrooke, 2500 Boulevard de l'Université,  
17   Sherbrooke QC J1K 2R1, Canada; <sup>i</sup> Department of Biological Sciences, University of Toronto Scarborough,  
18   Toronto ON M1C 1A4, Canada; <sup>j</sup> Department of Integrative Biology, University of Guelph, 50 Stone Road  
19   East, Guelph ON N1G 2W1, Canada; <sup>k</sup> Department of Geography, University of British Columbia, 2329  
20   West Mall, Vancouver BC V6T 1Z4, Canada; <sup>l</sup> Department of Biology, University of British Columbia, 1177  
21   Research Road, Kelowna BC V1V 1V7, Canada; <sup>m</sup> Department of Earth, Ocean, and Atmospheric Sciences,  
22   University of British Columbia, 2207 Main Mall, Vancouver BC V6T 1Z4, Canada

23

24

25   **Authors' email addresses:** jcurrie@wwfcanada.org (JC); joseph.burant@mcgill.ca (JBB);  
26   valentina.marconi@ioz.ac.uk (VM); blain@zoology.ubc.ca (SAB); emry@zoology.ubc.ca (SE) ;  
27   katherine.hebert@usherbrooke.ca (KH); garland.xie@mail.utoronto.ca (GX);  
28   nicole.moore@mail.mcgill.ca (NM); wangx@uoguelph.ca (XW); andrea.brown3@mail.mcgill.ca (AB);  
29   grevstad@student.ubc.ca (LG); louise.mcrae@ioz.ac.uk (LM); stefano.mezzini@ubc.ca (SM);  
30   p.pata@oceans.ubc.ca (PP); robin.freeman@ioz.ac.uk (RF)

31

32   **Corresponding author:** jcurrie@wwfcanada.org

33    **Abstract**

34    To effectively combat the biodiversity crisis, we need ambitious targets and reliable indicators to  
35    accurately track trends and measure conservation impact. In Canada, the Living Planet Index (LPI) has  
36    been adapted to produce a national indicator by both WWF-Canada (Canadian Living Planet Index; C-LPI)  
37    and Environment and Climate Change Canada (Canadian Species Index) to provide insight into the status  
38    of Canadian wildlife, by evaluating temporal trends in vertebrate population abundance. The indicator  
39    includes data for just over 50% of Canadian vertebrate species. To assess whether the current dataset is  
40    representative of the distribution of life history characteristics of Canadian wildlife, we analyzed the  
41    representation of species-specific biotic variables (i.e., body size, trophic level, lifespan) for vertebrates  
42    within the Canadian-LPI compared to native vertebrates lacking LPI data. Generally, there was  
43    considerable overlap in the distribution of biotic variables for species in the C-LPI compared to native  
44    Canadian vertebrate species lacking LPI data. Statistically, however, significant differences among  
45    distributions were found, driven in large part by discrepancy in the representation of fishes — where the  
46    C-LPI included larger-bodied and longer-lived species. We provide recommendations for targeted data  
47    collection and additional analyses to further strengthen the applicability, accuracy and representativity  
48    of biodiversity indicators.

49

50    **Key Words:** Living Planet Index, representation, biodiversity indicator, Canada, conservation

51

52

53

54

55

56

57 **Introduction**

58 Combatting the biodiversity crisis requires ambitious targets, widespread action, and reliable indicators  
59 to accurately track the impact of conservation intervention and progress towards national (e.g.,  
60 Canada's Federal Sustainable Development plan) and international (e.g., Convention on Biological  
61 Diversity) goals. The Living Planet Index (LPI) — conceived nearly 25 years ago — is a biodiversity  
62 indicator used to evaluate the state of wildlife by assessing temporal trends in vertebrate population  
63 abundance. The indicator has been applied at global (WWF 2020), regional (see for example McRae *et*  
64 *al.* 2012) and national (Bayraktarov *et al.* 2020, Marconi *et al.* 2021) geographic scales and used as a  
65 public engagement and advocacy tool to enhance ambition for wildlife conservation (Ledger *et al.* 2022;  
66 in preparation).

67

68 In Canada, the LPI methodology has been adapted to produce a national indicator by both WWF-Canada  
69 (Canadian Living Planet Index; C-LPI) (WWF-Canada 2020) and Environment and Climate Change Canada  
70 (Canadian Species Index; CSI) (ECCC 2019). These indicators reveal near-stable trends of vertebrate  
71 population abundance since 1970, on average (+6% and -4%, respectively, owing to differences among  
72 data selection). The dataset underlying the C-LPI, largely publicly available via the Living Planet Index  
73 database ([www.livingplanetindex.org](http://www.livingplanetindex.org)) and continuously augmented, now incorporates thousands of  
74 time series from more than 900 vertebrate species, representing 50.9% of native vertebrate species in  
75 Canada (CESCC 2016). The taxonomic and spatial representation of the dataset behind the indicator has  
76 been analyzed in Marconi *et al.* (2021) which showed that Canadian birds are better represented in  
77 comparison to other taxonomic groups. However, it's also important to know whether species currently  
78 included in the C-LPI database are representative of the distributions of biological and ecological traits  
79 observed in the broader set of vertebrate species that occur in Canada.

80

81 The biases in our data broadly reflect the biases in biodiversity research more generally (Donaldson *et*  
82 *al.* 2016, Troudet *et al.* 2017). Monitoring is important (Tittensor *et al.* 2014) and additional data  
83 collection is key to establishing accurate trends in currently monitored species. However, it seems  
84 unlikely that there will be sufficient data to provide us with an unbiased picture of biodiversity trends  
85 overall, at least in the short-term. In the meantime, intrinsic (biological) and extrinsic (ecological) traits  
86 may provide information on species' vulnerability to anthropogenic threats and therefore on their  
87 predicted population trends and risk of extinction (see Chichorro *et al.* 2019 for a comprehensive  
88 review). Without insight into some of the characteristics of the species that are not currently monitored  
89 (i.e., insight into data gaps), however, we may be wrongly estimating the direction and magnitude of  
90 temporal trends in population abundance.

91

92 Here we analyze the representation of species-specific biotic variables for vertebrates within the C-LPI  
93 compared to the broader vertebrate subphylum in Canada to assess whether the current dataset is  
94 representative of the distribution of life history characteristics of Canadian wildlife. In addition to  
95 assessing the distribution of common biotic variables among taxa, we collected data for a variety of  
96 supplementary biotic and abiotic variables that we are making publicly available for future analyses.

97

## 98 **Methods**

### 99 ***Species list: Canadian Living Planet Index***

100 We used the data underlying the Canadian-Living Planet Index (WWF-Canada 2020). These data are also  
101 included within the global LPI Data Portal ([www.livingplanetindex.org/data\\_portal](http://www.livingplanetindex.org/data_portal)) (with the exception  
102 of records that were provided to the LPI team under the agreement they would be kept confidential;  
103 e.g., data for sensitive populations or at-risk species) — a repository that contains thousands of records  
104 depicting temporal changes in vertebrate abundance. The data have been collated from a variety of

105 sources, including peer-reviewed publications, government databases, and grey literature. The criteria  
106 for inclusion within the dataset are as follows: a) the data have been collected using comparable  
107 methods for at least two years for the same population (aligns with the global LPI and CSI methodology;  
108 the C-LPI methodology has recently been improved to limit inclusion for populations that have at least  
109 three years of data collection); b) data are specific to units of population size, either a direct measure  
110 such as population counts, densities, or indices, or a reliable proxy such as breeding pairs, nests, tracks,  
111 capture per unit effort or measures of biomass for a single species; and c) the source is referenced and  
112 traceable (Collen *et al.* 2009). Species classified as “presumed extirpated”, “probably extirpated” or “not  
113 applicable” by the Wild Species Report (CESCC 2016) have been excluded from the dataset. The latter  
114 category includes exotic species, hybrids, or species occurring infrequently and unpredictably in Canada.  
115 Non-native species are excluded from the dataset. These criteria resulted in a final subset of 906 species  
116 that commonly occur in Canada.

117

118 ***Species list: Native Canadian vertebrates***

119 Species listed within the 2015 Wild Species Reports (CESCC 2016) were compiled into a complementary  
120 dataset to compare traits of species within the C-LPI, to the broader group of native vertebrates found in  
121 Canada (C-Vertebrates) and those lacking LPI data (C-Vertebrates Only). We included native Canadian  
122 species with applicable conservation status (i.e., exotic species, hybrids and accidental species under the  
123 NatureServe rank of “not applicable” were excluded). Discrepancies in nomenclature between the C-LPI  
124 taxonomic authorities and Wild Species Reports (CESCC 2016) necessitated evaluation of 58 species,  
125 which were resolved for comparison using the *taxize* R package (Chamberlain *et al.*, 2013). For some  
126 species, we matched those that had synonyms for the binomial Latin names; others were resolved to  
127 address discrepancies between North American (greater species specificity) and global (where species  
128 are sometimes lumped together) nomenclature. In total, there are 1,779 extant, native vertebrate

129 species in Canada — more than half of which ( $n = 906$ ) have population trend data within the broader C-  
130 LPI dataset.

131

132 ***Traits***

133 Abiotic and biotic species trait data (i.e., life history characteristics, habitats and threats) were extracted  
134 from a variety of publicly accessible online databases for birds (Table S1), fish (Table S2), mammals  
135 (Table S3), amphibians and reptiles (Table S4). Data extraction was conducted separately for each taxon,  
136 given differences among common variables and units reported across taxonomic groups. For instance,  
137 length was typically recorded for fish and herpetofauna (Froese and Pauly 2016, Santini *et al.* 2018),  
138 while mass was more commonly reported for birds and mammals (Myhrvold *et al.* 2015). For each  
139 taxonomic group, variables with sufficient data were mapped onto species from the C-LPI, creating  
140 multi-variable datasets that are publicly available for future analyses.

141

142 ***Biotic variables***

143 *i. Birds*

144 Six biotic variables were selected for inclusion within the C-LPI Bird Trait database (Table S1). Adult body  
145 mass, lifespan (maximum longevity) and mean longevity were extracted from the Amniote database  
146 (Myhrvold *et al.* 2015), using the R package *traitdata* (RS-eco 2021). Data on body mass was also  
147 available from the EltonTraits database (Wilman *et al.* 2014), but the Amniote database (Myhrvold *et al.*  
148 2015) was selected as the primary source given the quality of its metadata (i.e., more extensive).  
149 Average values per species were calculated for each biotic variable. In addition, hand-wing index (i.e., a  
150 measure of wing aspect ratio and a proxy for dispersal ability) and dietary guild were extracted from the  
151 Global Hand-Wing Index repository (Sheard *et al.* 2020). Dietary guild categories were aligned to trophic

152 level categorization (carnivore, omnivore and herbivore). One species (turkey vulture, *Cathartes aura*)  
153 was assigned to the carnivore category as it almost exclusively feeds on carrion.

154

155 *ii. Fishes*

156 Data were extracted from FishBase (Froese and Pauly 2016) using the R package *rfishbase* (Boettiger *et*  
157 *al.* 2012). Fourteen biotic variables were selected for inclusion within the C-LPI Fishes Trait database  
158 (Table S2), including length, weight and lifespan. Average values per species were calculated for each  
159 biotic variable.

160

161 *iii. Mammals*

162 Gestation period and lifespan (maximum longevity) were retrieved from the Amniote database  
163 (Myhrvold *et al.* 2015) using the R package *traitdata* (RS-eco 2021), while body size and diet were  
164 extracted from EltonTraits (Wilman *et al.* 2014). In total, four biotic variables were selected for inclusion  
165 within the C-LPI Mammal Trait database (Table S3). Mammal trophic level was calculated using diet  
166 composition, where herbivore was defined as species associated with an entirely vegetarian diet (e.g.,  
167 plants, seeds, nectar and fruit), carnivores were classified as species that consume other animals, and  
168 omnivores included species with mixed diets.

169

170 *iv. Amphibians & Reptiles*

171 Seven biotic variables were selected for inclusion within the C-LPI Amphibian and Reptile Trait database  
172 (Table S4), including variables related to body size, diet and reproduction. Data were extracted from  
173 AmphiBIO (Oliveira *et al.* 2017), the Amniote database (Myhrvold *et al.* 2015), Meiri 2018, Santini *et al.*  
174 2017, Atwood *et al.* 2020, and SquamataBase (Grubler 2020) using *traitdata* (RS-eco 2021) and  
175 *squamatabase* (Grubler 2020) R packages. Trait data that were comparable across datasets (e.g., body

176 mass) were hierarchically extracted. For example, body mass was first extracted from AmphiBIO  
177 (Oliveira *et al.* 2017), then from the amniote life history database (Myhrvold *et al.* 2015) for species that  
178 lacked data in AmphiBIO. Because multiple entries were available per species within the amphibian  
179 allometry database (Santini *et al.* 2017) values with the highest sample sizes were selected for inclusion.  
180 Amphibians and reptiles were grouped together as herpetofauna to improve sample size for analysis.

181

182 ***Abiotic variables***

183 Geographic ranges of species tagged with distributions in Canada were retrieved from the International  
184 Union for Conservation of Nature (IUCN) Red List (2021; version 2021-2). Ranges were clipped to Canada  
185 and its associated Exclusive Economic Zone to account for the presence of marine species. The  
186 proportion of the global range in Canada was also calculated. For each species range in Canada, mean  
187 topographic and climate variables were extracted, representing coarse characteristics of the habitat  
188 (Table S5). Environmental variables were derived from the 1961-1990 reference period, which partially  
189 overlaps with the temporal scale of the C-LPI. Similarly, habitat descriptions and IUCN status for species  
190 were extracted from the IUCN Red List using the *rredlist* package. Due to the hierarchical nature of  
191 habitat associations, the data is summarized into the number of habitats for Levels I and II of the *Habitat*  
192 *Classification Scheme v3.1* (IUCN 2012). Only species with associated IUCN data were included, resulting  
193 in a smaller sample for any subsequent analyses.

194

195 All data and associated code for this analysis are available [online](#).

196

197 ***Analyzing the representation of biotic variables***

198 Three biotic variables were selected from the taxon-specific datasets to create a more fulsome cross-  
199 taxa database. Selected variables included relative body size, trophic level and lifespan. Continuous

200 variables of body size and lifespan were log-transformed, while trophic level included three categorical  
201 levels (herbivore, omnivore, carnivore). The overarching cross-taxa C-LPI Trait Database included 836  
202 species, while the broader C-Vertebrate Trait Database included 1679 species.

203

204 Shapiro-Wilk tests were used to assess the distributions of our continuous variables of body size and  
205 lifespan for species included within the C-LPI trait database and for other native Canadian vertebrate  
206 species lacking LPI data (C-Vertebrates Only) (Figure S1). We assessed the percent overlap among  
207 distributions for continuous variables and applied two-sample nonparametric Kolmogorov-Smirnov (KS)  
208 tests to evaluate differences between the distribution of traits for species in the C-LPI Trait database and  
209 the C-Vertebrate (Only) Trait database. We also assessed the difference in categorical trophic level  
210 distribution using Pearson's and Fisher's Exact chi-squared tests. All tests were performed for each biotic  
211 variable at the taxonomic level (i.e., birds, mammals, fish, amphibians and reptiles), and with all species  
212 combined. Completeness of variables in our dataset ranged between 35.0 and 87.8%. In order to rerun  
213 the analysis with a more complete dataset, we imputed missing values via nonparametric imputation  
214 using Random Forest (Stekhoven and Buehlmann 2012).

215

216 To supplement our analysis of the two continuous traits (body size and lifespan), we applied a  
217 nonparametric bootstrapping approach to generate 95% confidence (95% CIs) on the median difference  
218 between datasets (i.e., the effect size; Cohen 1994; Ho *et al.* 2019). Bootstrapping was done via  
219 randomized resampling with replacement ( $n = 5,000$  iterations) of the C-LPI and C-Vertebrate (Only)  
220 subsets of species-specific trait values to generate resamples equal in size to the original subsets. In  
221 addition, bootstrapping was performed twice: once using an unstratified approach in which all species  
222 were treated as a single sample without consideration for taxonomic representation, and once using  
223 taxon-specific resampling in which data were sampled proportionally for each of the four vertebrate

224 groups under consideration. We include the median and corresponding 95% CI on the difference  
225 between subsets for each of trait comparisons as a data-centered complement to the statistical results,  
226 and present Gardner-Altman estimate plots of these effects in the Supplementary Materials.

227

228 Analyses and visualization were conducted using R statistical software (v. 4.1.1; R Core Team 2021),  
229 making use of the *missforest* (Stekhoven 2013), *tidyverse* (Wickham *et al.* 2019), *dabestr* (Ho *et al.* 2019),  
230 *ggpubr* (Kassambara 2020), *effectsize* (Ben-Schachar *et al.* 2020), *data.table* (Dowle and Srinivasan  
231 2021), *traitdata* (RS-eco 2021), and packages.

232

## 233 **Results**

### 234 ***Dataset description***

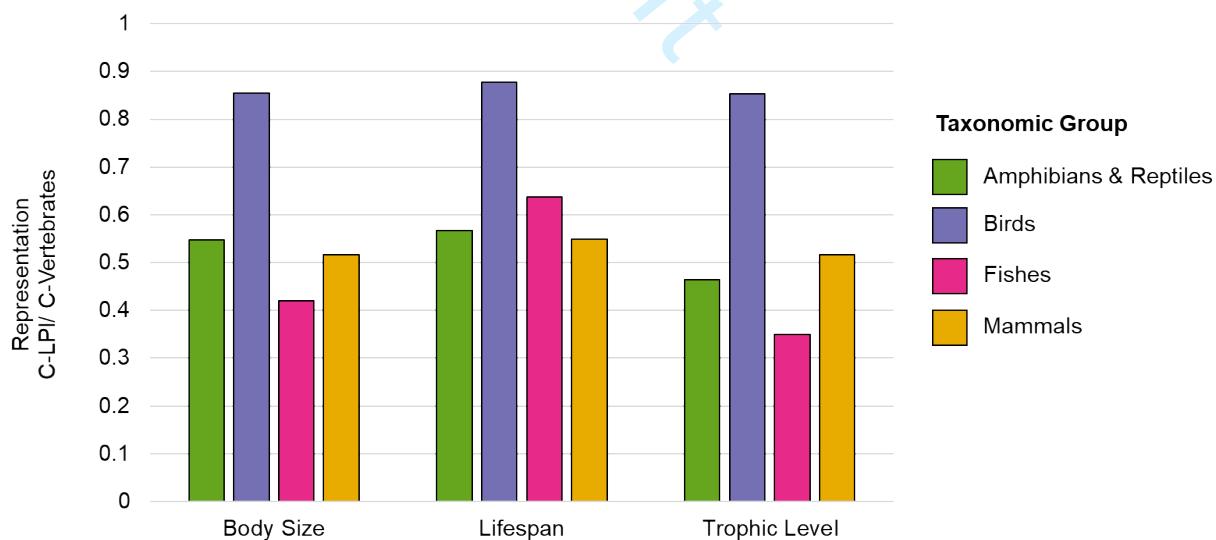
235 While the C-LPI contains data for 50.9% (906/1779) of native vertebrate species in Canada, not all of  
236 these species had available trait data (Table 1). In total, 49.8% (836/1679) of Canadian vertebrates with  
237 data for at least one biotic variable were included within the C-LPI Trait database, resulting in a trait  
238 database covering 92.3% of species (836 of 906) included in the C-LPI. Overall, birds tended to have the  
239 best data coverage in terms of both population time series and traits, with over 85% of species with trait  
240 data included within the C-LPI Trait database (Figure 1). Birds were the best represented, while the  
241 representation of fish trait data was more variable; only 35.0% (352/1007) of fishes with available trait  
242 information were included within the C-LPI Trait database relative to the broader C-Vertebrates Trait  
243 database. In total, 629 species within the C-LPI Trait database, and 883 native Canadian vertebrate  
244 species lacking LPI data (C-Vertebrates Only) contained trait data for all three biotic variables.  
245 Impressively, 94.4% (1679/1779) of native Canadian vertebrates were included in the analysis.

246

247 Table 1. Data availability for assessing the representation of species included within the C-LPI, including  
 248 proportion of vertebrates within Canada, the relative representation of biotic traits.

Data Availability	Birds	Fish	Herps	Mammals	All
C-LPI (Representation in C-LPI)	388 (42.8%)	370 (40.8%)	47 (5.2%)	101 (11.1%)	906
C-Vertebrates (Representation in C-Vertebrates)	453 (25.5%)	1044 (58.7%)	88 (4.9%)	194 (10.9%)	1779
<b>Species representation (%)</b>	<b>85.7</b>	<b>35.4</b>	<b>53.4</b>	<b>52.1</b>	<b>50.9</b>
C-LPI Trait Database (Representation in C-LPI Trait Database)	344 (41.1%)	352 (42.1%)	46 (5.5%)	94 (11.2%)	836
C-Vertebrates Trait Database (Representation in C-Vertebrates Trait Database)	404 (24.1%)	1007 (60.0%)	86 (5.1%)	182 (10.8%)	1679
<b>Species trait representation (%)</b>	<b>85.1</b>	<b>35.0</b>	<b>53.5</b>	<b>51.6</b>	<b>49.8</b>
C-Vertebrates Only Trait Database (Representation in C-Vertebrates Only Trait Database)	60 (7.1%)	655 (77.7%)	40 (4.7%)	88 (10.4%)	843

249



250  
 251 Figure 1. Data availability for assessing the distribution of species traits. Bars depict the representation  
 252 of species with trait data in the C-LPI, relative to the number of species with trait data for the broader  
 253 vertebrate subphylum in Canada (C-Vertebrates).

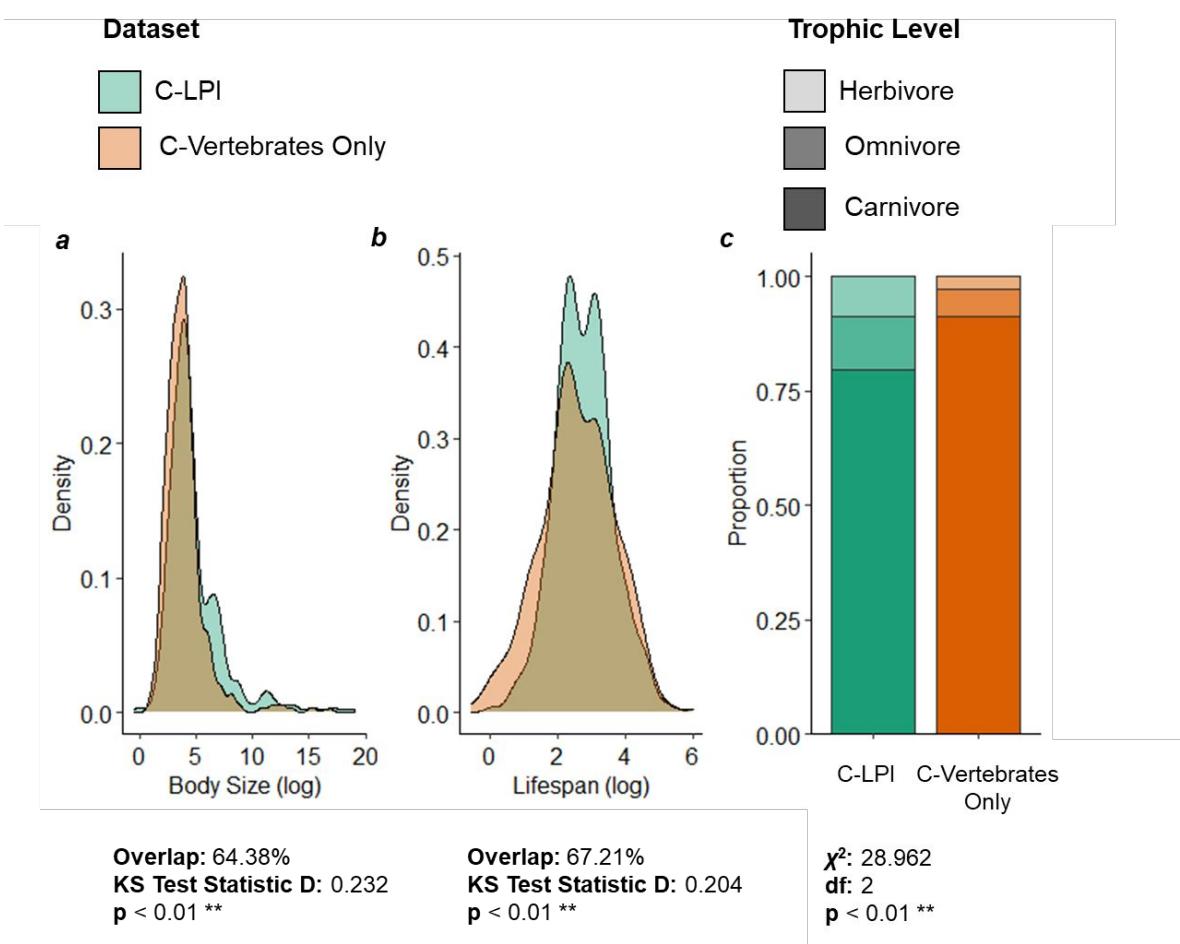
254 ***Representation of biotic variables***

255 The distribution of body size, lifespan and trophic level for species within the C-LPI were largely  
256 overlapping (> 64%) with the distribution of biotic variables for native Canadian vertebrate species  
257 lacking LPI data (C-Vertebrates Only) (Figure 2; Table S6) and the broader vertebrate subphylum (>82%)  
258 nationally (i.e., including LPI species) (Table S7; Figure S7). Despite the broad overlap, statistically,  
259 significant differences between distributions were found for body size (KS test statistic D: 0.23, p < 0.01),  
260 lifespan (KS test statistic D: 0.20, p < 0.01) and trophic level ( $\chi^2$ : 28.96, df:2, p < 0.01), largely due to  
261 differences among fish distributions (Table S6) — the taxonomic group that contributed the most data  
262 yet had the poorest overlap in distributions (< 54%). Overall, there were a greater number of fishes with  
263 large body sizes (Figure 3; average log body size: 3.93; average body size: 50.91) and long lifespans  
264 (Figure 4; average log lifespan: 2.93; average lifespan: 18.73) compared to species lacking LPI data  
265 (average log body size: 3.56; average body size: 35.12; average log lifespan: 2.58; average lifespan:  
266 13.20). Birds and mammals also had statistically significant distributions for body size and lifespan  
267 among datasets but exhibited larger overlap (>54%) among distributions in comparison to fish. Although  
268 a significant difference among trophic levels was detected when taxonomic groups were combined  
269 (Table S6), statistical differences were not detected within each taxonomic group (Figure 5). In addition,  
270 relationships among biotic variables were similar for species within the C-LPI and those lacking LPI data  
271 (C-Vertebrates Only) and appeared to be correlated (Figure S2).

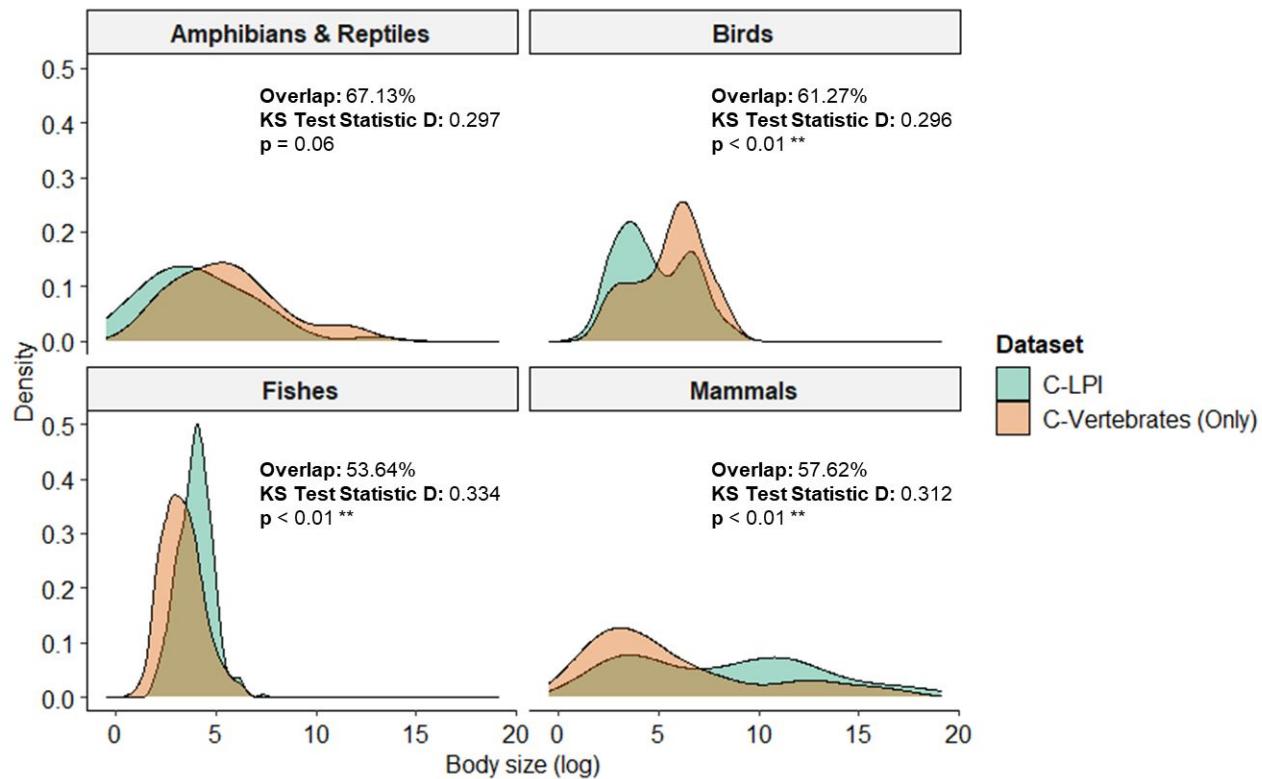
272

273 Although the distribution of biotic variables for species within the C-LPI generally overlapped with  
274 species lacking LPI data (C-Vertebrates Only), we were able to pinpoint traits within taxonomic groups  
275 that may help to improve coverage across Canada. For instance, the C-LPI should work to incorporate  
276 temporal population estimates data for large birds and small fishes (Figure 3) — which correlate to long-  
277 lived birds and fish with short lifespans (Figure 4) — in addition to carnivorous amphibians and reptiles

278 (Figure 5). These results are reinforced through a bootstrapping approach (i.e., repeated resampling  
 279 with replacement to generate samples of the same size as the reference) that depicted considerable  
 280 overlap between the bootstrapped datasets, yet significant differences among distributions were  
 281 detected when assessing all species combined for body size (median difference between groups = -0.72,  
 282 95% CI = -0.87, -0.56; Figure S3) and lifespan (median difference between groups = -0.45, 95% CI = -0.64,  
 283 -0.27; Figure S4). Likewise, significant differences among distributions were also apparent when  
 284 individual taxonomic groups, with the exception of amphibians and reptiles (Figure S5 and Figure S6).

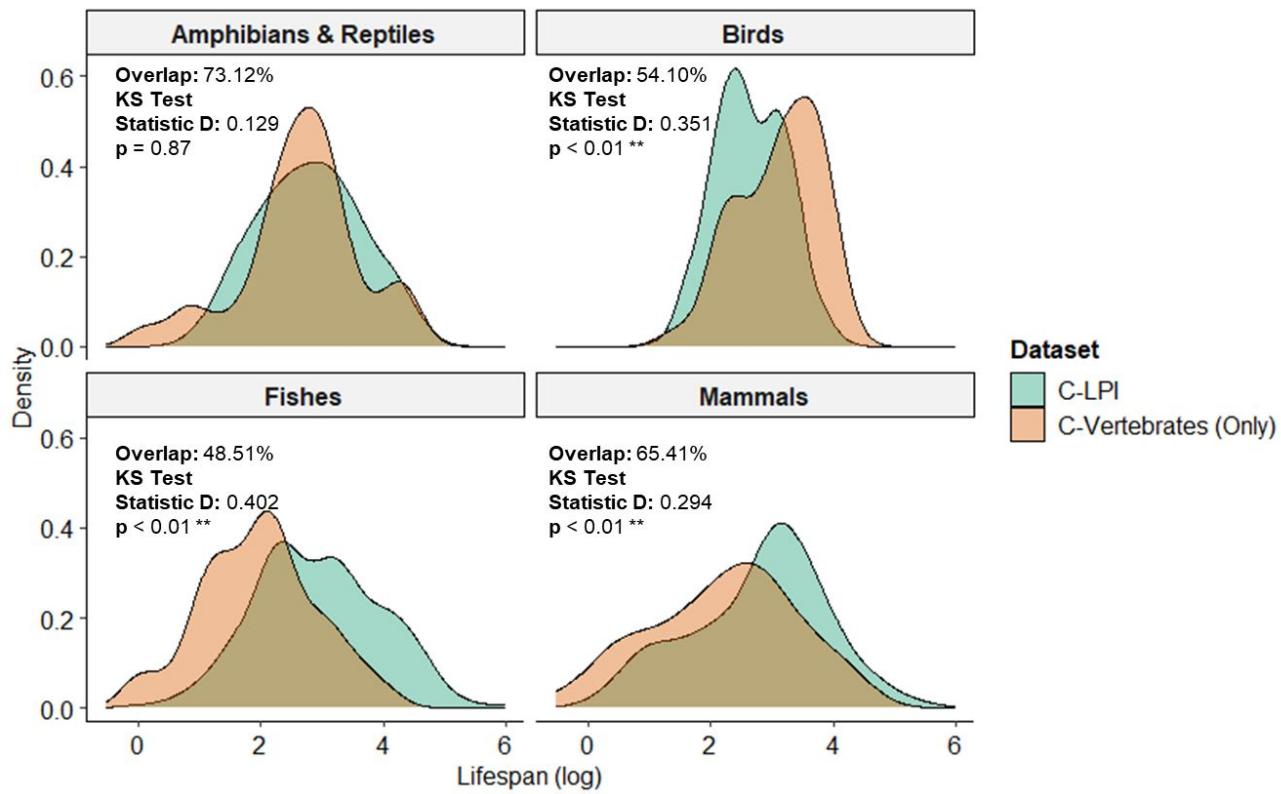


286 Figure 2. Comparing distributions of (a) body size, (b) lifespan, and (c) trophic level for species included  
 287 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
 288 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.



289

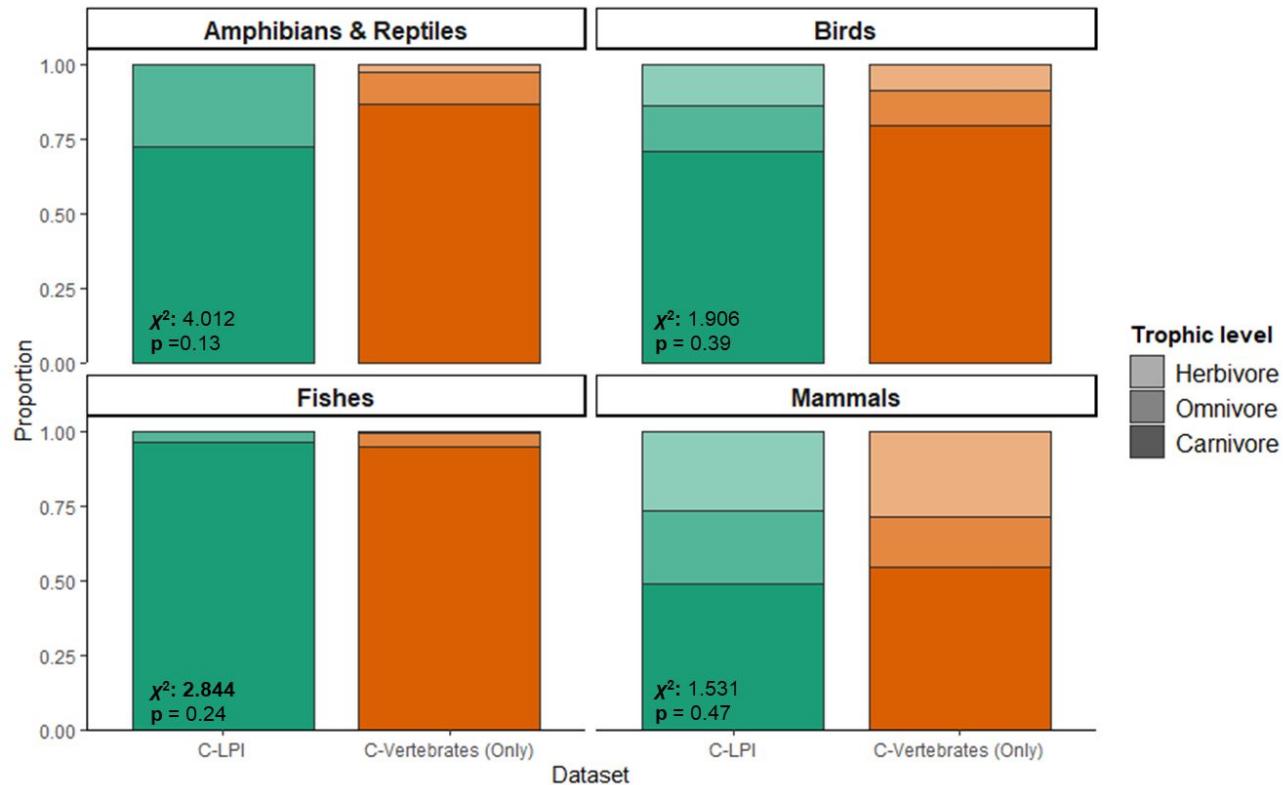
290 Figure 3. Comparing distributions of bird, fish, herpetofauna, and mammal body size for species included  
291 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
292 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.



293

294 Figure 4. Comparing distributions of bird, fish, herpetofauna, and mammal lifespan for species included  
 295 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
 296 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

297



298

299 Figure 5. Comparing distributions of bird, fish, herpetofauna, and mammal trophic level for species  
 300 included within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking  
 301 LPI data (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

302

303

304 Fish also had the poorest overlap in biotic traits when comparing the C-LPI to the broader vertebrate  
 305 subphylum nationally (Table S7). Conversely, the overlap in body size (Figure S8) and lifespan (Figure S9)  
 306 for birds improved largely due to bird data representation (i.e., sample size among datasets).

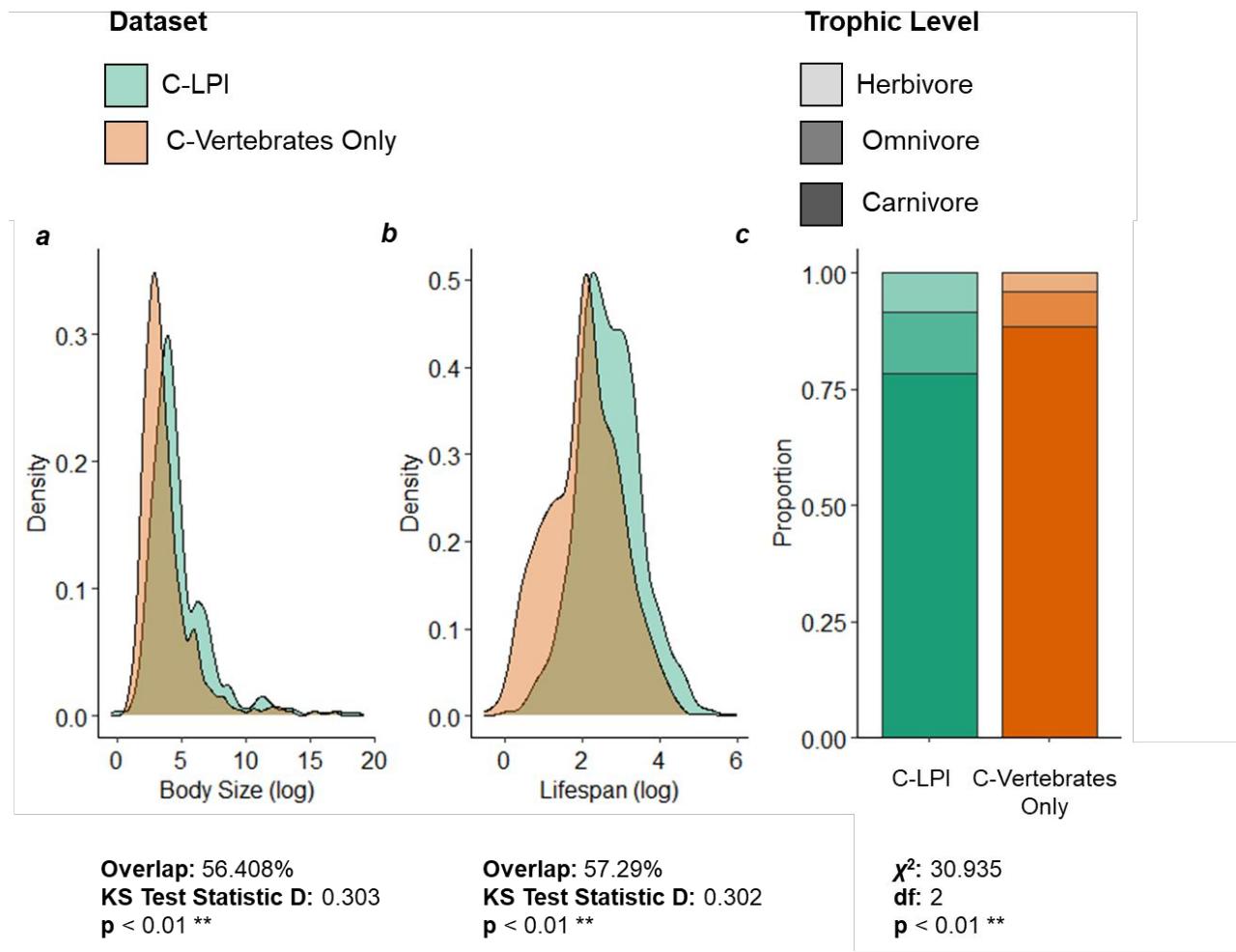
307 Interestingly, trophic level representation among species in the C-LPI relative to the broader vertebrate  
 308 subphylum nationally was nearly identical for birds, fishes and mammals (Figure S10).

309

310 When imputing values to increase dataset coverage, the number of species in the C-LPI trait database  
311 grew from 836 to 850 species, while the number of species in the C-Vertebrates Only database grew  
312 from 843 to 852 species. Consequently, only a handful of additional species were added to the analysis,  
313 but the data coverage by biotic variable increased dramatically — particularly for fish (Figure S11). When  
314 using imputed data to enhance coverage of biotic variables, the distribution of body size, lifespan and  
315 trophic level for species within the C-LPI exhibited slightly less overlap (>56% using imputed data  
316 compared to >64% for the original database) with the distribution of biotic variables for native Canadian  
317 vertebrate species lacking LPI data (C-Vertebrates Only) (Figure 6). Statistically significant differences  
318 among distributions were found for all three biotic variables and despite improved data coverage, fish  
319 still exhibited the smallest overlap in distributions relative to other taxonomic groups.

320

raft



321

322

323 Figure 6. Comparing distributions of (a) body size, (b) lifespan, and (c) trophic level for species included  
324 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
325 (C-Vertebrates Only) using imputed data. Asterisks (\*\*) denote statistical significance.

326

327

328

329

330

331

332 **Discussion**

333 ***Dataset description***

334 The analysis presented here included at least some data for over 95% of native Canadian vertebrates,  
335 boasting impressive data coverage. However, data coverage for individual traits varied. In these  
336 instances, taxonomic biases known to exist within published biodiversity data may have been inherited  
337 within the compiled datasets for this analysis — including, for example, threatened species (Donaldson  
338 *et al.* 2016) or those with public appeal (e.g., charismatic species) (Troudet *et al.*, 2017). Birds were the  
339 best represented taxonomic group, for each of the three biotic variables considered, largely attributable  
340 to the high representation of this taxonomic group within the dataset itself (Marconi *et al.* 2021). By  
341 contrast, the representation of trait data was more variable for fishes, where species with publicly  
342 available trait data ranged from 322 (of 1044 native species) for lifespan, to 792 for body size and 1018  
343 for trophic level. The availability of this information for Canadian species is consistent with data  
344 availability for other studies using Fishbase, which found fewer entries of lifespan and mortality, yet  
345 more data for growth parameters (Thorson *et al.* 2014). Importantly, while there now exists a plethora  
346 of publicly available information regarding species-specific life history characteristics on a global scale  
347 (e.g., Myhrvold *et al.* 2015, Froese and Pauly 2016), the information is not always accurate or precise  
348 (Thorson *et al.* 2014), and thus any subsequent analyses, such as the one presented here, would inherit  
349 the biases within the datasets used. Similarly, while the imputed dataset helped to enhance biotic trait  
350 coverage, the overlap in distributions declined compared to the non-imputed dataset, and it's possible  
351 that biases were introduced, affecting the results of the analysis as a consequence.

352

353 ***Representation of biotic variables***

354 Effective evidence-based wildlife management requires that conclusions and conservation decisions are  
355 informed by analyses of representative datasets and, where the data are not fully representative,

356 recognition that such underlying biases may impact conservation approaches. Consequently, the  
357 interpretation of indicators is inherently tied to considerations of data representation. In the context of  
358 the C-LPI, the species and population time series underlying the index must be representative of the  
359 diversity of Canadian vertebrates to be an effective biodiversity indicator. Yet, despite boasting some of  
360 the best representation of species included within a national-scale abundance indicator, trends for  
361 approximately half of native Canadian vertebrates are lacking within the current dataset (Marconi *et al.*  
362 2021). By evaluating and integrating the distribution of biotic traits for species within the C-LPI,  
363 compared to native Canadian vertebrate species lacking LPI data, we can target data gaps and recognize  
364 underlying data biases for more informed reporting of average trends in population abundance nation-  
365 wide.

366

367 While the distributions of biotic variables from species within the C-LPI overlap with the distributions of  
368 those for native Canadian vertebrate species lacking LPI data, some notable statistical discrepancies  
369 exist. For instance, the body size of fishes within the C-LPI skewed larger relative to the broader group of  
370 fishes in Canada. Most temporal population trend data for fishes were derived from marine Research  
371 Vessel Trawl Surveys conducted by Fisheries and Oceans (DFO) Canada (e.g., DFO 2020a, DFO 2020b,  
372 DFO 2020c), resulting in a bias towards marine over freshwater species and towards commercially  
373 important species. Although DFO trawl data date back to the 1940s, stratified random sampling for  
374 temporal evaluation of biological information was introduced in the 1970s, resulting in substantial  
375 temporal data coverage within Canada's Exclusive Economic Zone (Smith and Somerton 1981). Annual  
376 Research Vessel surveys on both Pacific and Atlantic coasts are conducted using trawls (i.e.,  
377 standardized equipment for comparison), that are more tailored to sampling groundfishes, which  
378 include the greatest number and highest proportion of stocks classified as critical (ECCC 2021).  
379 Moreover, their primary purpose is to gain species-specific information as a critical component for

380 science-based stock assessments for proper management of marine fisheries — where larger fishes are  
381 often targeted (Pauly *et al.* 2002). Consequently, this bias noted within the C-LPI is not particularly  
382 surprising. Moreover, while this bias exists, it is in many ways important that the data in body size skews  
383 towards large-bodied marine fishes compared to small-bodied species, as Canadian fishes are frequently  
384 threatened with overexploitation (Currie and Marconi 2020), less likely to receive legal protections  
385 under the Species at Risk Act (Mooers *et al.* 2007, Findlay *et al.* 2009) and large fish are under greater  
386 threat of global extinction (Olden *et al.* 2007).

387

388 While there was large overlap in biotic traits between datasets, our results indicate that monitoring and  
389 data collection should target temporal abundance trends for small fishes and mammals, and large birds,  
390 — which are correlated with short and long lifespans, respectively. In addition, efforts should be made  
391 to enhance the representation of carnivorous herpetofauna, particularly reptiles, which represent 70%  
392 of carnivorous herpetofauna not included within the C-LPI. Globally, abundance studies for reptiles are  
393 comparatively scarce (Saha *et al.* 2018). This additional data collection would further enhance the  
394 representation of biotic traits within the C-LPI, ensuring that it is as representative as possible, while  
395 recognizing that temporal population trend data likely does not exist for all vertebrate species in  
396 Canada. The number of species within the C-LPI doubled over the decade between the first national  
397 Living Planet Report Canada in 2007 (WWF-Canada *et al.* 2007) and the landmark update in 2017 (WWF-  
398 Canada 2017), owing to dedicated data collection efforts in recent years (Marconi *et al.* 2021). However,  
399 despite concerted efforts since, few additional species with publicly available temporal abundance data  
400 have been added to the C-LPI database (WWF-Canada 2020). Thus, dedicated monitoring programs may  
401 be required for missing species and specific biotic traits to further enhance representation in the  
402 absence of publicly available abundance data.

403

404 This work builds upon the taxonomic and spatial representation of the dataset analyzed in Marconi *et al.*  
405 (2021), which has been used to assess threats to at-risk vertebrates in Canada and potential  
406 management actions for their recovery (Currie and Marconi 2020). Moreover, variations on the  
407 approach outlined in this analysis could be useful for evaluating a representative sample of vertebrates  
408 needed for calculating an LPI at national and/or regional scales beyond Canada — a frequent question  
409 that has yet to be remedied, though studies are underway (Pata *et al.* 2020 in preparation).

410

#### 411 **Moving forward**

412 The alignment of biotic traits uncovered through the analysis helps to reinforce the value of the C-LPI at  
413 a national level — including its use as an ecological indicator to track progress towards biodiversity  
414 goals, and as a tool for communication and policy. Nevertheless, there are a variety of additional  
415 considerations, analysis and data collection needed to further strengthen the applicability, accuracy and  
416 precision of the indicator. For instance, data monitoring and collection could target small fishes and  
417 mammals, and large birds (correlating to short and long lifespans, respectively, in addition to  
418 carnivorous reptiles. The LPI (and derived indicators such as the C-LPI and CSI) are meant to reveal high-  
419 level, long-term trends in biodiversity. While full species representation may not be possible, targeted  
420 data monitoring and collection may serve to strengthen the indicator and its use for policy and  
421 communication.

422

423 From an indicator perspective, the C-LPI (and CSI) and current analysis are specific to native vertebrate  
424 species, which represent a small proportion of the 80,000 known species in Canada (CESCC 2016) and  
425 thus a more fulsome account of biodiversity trends may be needed to detect changes at multiple  
426 taxonomic and temporal levels. Similarly, while enhanced representation of species strengthens the  
427 accuracy of the indicator, improving the geographic coverage of data collection to ensure that

428 population trends wholly reflect patterns at a national scale will be critical in addressing geographic  
429 biases. Buschke *et al.* (2021) highlight the usefulness of randomization null models for improving  
430 accuracy and presenting uncertainty around the LPI— which may be useful for overcoming random  
431 population fluctuations that bias quantitative estimates. Leung *et al.* (2020) propose investigating  
432 clusters of extreme decline or increase from the LPI to inform decision-makers and prioritize  
433 conservation efforts, as averages tend to mask important trends. Yet moving beyond the use of the LPI  
434 as a biodiversity indicator to action, bending the curve of biodiversity loss (Mace *et al.* 2018) requires an  
435 integrated strategy underpinned by immediate effort and unprecedented ambition (Leclère *et al.* 2020).

436

#### 437 **Acknowledgements**

438 The manuscript is a product of a virtual working group supported by the Living Data Project (LDP), an  
439 initiative of the Canadian Institute of Ecology and Evolution. We are grateful to the LDP, led by Professor  
440 Diane Srivastava (UBC), for providing the financial support and for bringing together a wonderful group  
441 of students to collaboratively develop a trait database for species included in the Canadian Living Planet  
442 Index. The LDP is funded by a Collaborative Research and Training Experience (CREATE) grant from the  
443 Natural Sciences and Engineering Research Council of Canada. We also thank Dr. Gracielle Higino for her  
444 help in organizing and facilitating the working group, and Francis Banville for his participation. We are  
445 grateful to Sally Otto and Xav Harrison for their statistical expertise and guidance in strengthening the  
446 manuscript. The data used in our analyses have been collected by countless researchers and field  
447 technicians, and so we thank them for sharing their hard-earned data and making this work possible. We  
448 acknowledge that Canada is colonized land; data have been collected on the traditional, ancestral,  
449 treaty, and unceded territories of many Indigenous peoples, often without consultation or consent.  
450 Reconciliation requires that we progress in a way that respects Indigenous connections to these lands,  
451 with a shared sense of collective responsibility for the wildlife that call them home.

452 **References**

453 Atwood TB, Valentine SA, Hammill E, McCauley DJ, Madin EMP, Beard KH, *et al.* 2020. Herbivores at the  
454 highest risk of extinction among mammals, birds and reptiles. *Science Advances*, 6(32). DOI:

455 [10.1126/sciadv.abb8458](https://doi.org/10.1126/sciadv.abb8458).

456

457 Bayraktarov E, Ehmke G, Tulloch AIT, Chauvenet AL, Avery-Gomm S, *et al.* 2020. A threatened species  
458 index for Australian birds. *Conservation Science and Practice*, 3(2): e322. DOI: [10.1111/csp2.322](https://doi.org/10.1111/csp2.322).

459

460 Ben-Shachar MS, Lüdecke D, and Makowski D. 2020. effectsize: Estimation of effect size indices and  
461 standardizedpParameters. *Journal of Open Source Software*, 5(56): 2815. DOI: [10.21105/joss.02815](https://doi.org/10.21105/joss.02815).

462

463 Boettiger C, Lang D, and Wainwright PC. 2012. rfishbase: exploring, manipulating and visualizing  
464 FishBase data from R. *Journal of Fish Biology*, 81(6): 2030-2039. DOI: [10.1111/j.1095-8649.2012.03464.x](https://doi.org/10.1111/j.1095-8649.2012.03464.x).

465

466 Buschke FT, Hagan JG, Santini L, and Coetzee BWT. 2021. Random population fluctuations bias the Living  
467 Planet Index. *Nature Ecology & Evolution*, 5: 1145-1152. DOI: [10.1038/s41559-021-01494-0](https://doi.org/10.1038/s41559-021-01494-0).

468

469 CESCC. 2016. Wild Species 2015: The General Status of Species in Canada. *Canadian Endangered Species  
470 Conservation Council*. N.G.S.W. Group. [online]: Available from: <https://www.wildspecies.ca/reports>.

471

472 Chamberlain S, Szoechs E, Foster Z, Arendsee Z, Boettiger C, Ram K, *et al.* 2020. Taxize: Taxonomic  
473 information from around the web. R package version 0.9.98.

474

- 475 Chichorro F, Juslén A, and Cardoso P. 2019. A review of the relation between species traits and  
476 extinction risk. *Biological Conservation*, 237: 220-229. DOI: [10.1016/j.biocon.2019.07.001](https://doi.org/10.1016/j.biocon.2019.07.001).
- 477
- 478 Cohen J. 1994. The earth is round ( $p < .05$ ). *American Psychologist*, 49(12): 997-1003. DOI:  
479 [10.1037/0003-066X.49.12.997](https://doi.org/10.1037/0003-066X.49.12.997).
- 480
- 481 Collen B, Loh J, Whitmee S, McRae L, Amin R, and Baillie JEM. 2009. Monitoring change in vertebrate  
482 abundance: the Living Planet Index. *Conservation Biology*, 23(2): 317-327. DOI: [10.1111/j.1523-1739.2008.01117.x](https://doi.org/10.1111/j.1523-<br/>483 1739.2008.01117.x).
- 484
- 485 Currie J, and Marconi V. 2020. An analysis of threats and factors that predict trends in Canadian  
486 vertebrates designated as at-risk. *FACETS*, 5(1): 49-66. DOI: [10.1139/facets-2019-0017](https://doi.org/10.1139/facets-2019-0017).
- 487
- 488 DFO. 2020a. Maritimes spring research vessel survey. *Fisheries and Oceans Canada*. [online]: Available  
489 from: <https://open.canada.ca/data/en/dataset/fecf045a-95a2-4b69-8a40-818649a62716>.
- 490
- 491 DFOb. 2020. NAFO Division 4T groundfish research vessel trawl survey (September Survey) dataset.  
492 *Fisheries and Oceans Canada*. [online]: Available from:  
493 <https://open.canada.ca/data/en/dataset/fecf045a-95a2-4b69-8a40-818649a62716>.
- 494
- 495 DFOc. 2020. Strait of Georgia Synoptic Bottom Trawl Survey. *Fisheries and Oceans Canada*. [online]:  
496 Available from: <https://open.canada.ca/data/en/dataset/fecf045a-95a2-4b69-8a40-818649a62716>.
- 497

- 498 Donaldson MR, Burnett NL, Braun DC, Suski CD, Hinch SG, Cooke SJ, *et al.* 2016. Taxonomic bias and  
499 international biodiversity conservation research. *FACETS*, 1: 105-113. DOI: [10.1139/facets-2016-0011](https://doi.org/10.1139/facets-2016-0011).
- 500
- 501 Dowle M, and Srinivasan A. 2021. *data.table*: Extension of `data.frame` . R package version 1.14.2.  
502 [online]: Available from: <https://CRAN.R-project.org/package=data.table>.
- 503
- 504 ECCC. 2019. Canadian Environmental Sustainability Indicators: Canadian species index. *Environment and*  
505 *Climate Change Canada*. Gatineau, Quebec. [online]: Available from:  
506 <https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/canadian-species-index.html>.
- 508
- 509 ECCC. 2021. Canadian Environmental Sustainability Indicators: Status of major fish stocks. *Environment*  
510 *and Climate Change Canada*. Gatineau, Quebec. [online]: Available from:  
511 [https://www.canada.ca/content/dam/eccc/documents/pdf/cesindicators/status-major-fish-stocks/2021/status\\_major\\_fish\\_stock.pdf](https://www.canada.ca/content/dam/eccc/documents/pdf/cesindicators/status-major-fish-stocks/2021/status_major_fish_stock.pdf).
- 513
- 514 Findlay CS, Elgie S, Giles B, and Burr L. 2009. Species listing under Canada's Species at Risk Act.  
515 *Conservation Biology*, 23(6): 1609–1617. DOI: [10.1111/j.1523-1739.2009.01255.x](https://doi.org/10.1111/j.1523-1739.2009.01255.x).
- 516
- 517 Froese R, and Pauly D. 2016. FishBase [online]: Available from [www.fishbase.org](http://www.fishbase.org)
- 518
- 519 Grubler MC. 2020. SquamataBase: a natural history database and R package for comparative biology of  
520 snake feeding habits. *Biodiversity Data Journal*, 8: e49943. DOI: [10.3897/BDJ.8.e49943](https://doi.org/10.3897/BDJ.8.e49943).
- 521

- 522 Ho J, Tumkaya T, Aryal S, Choi H, and Claridge-Chang A. 2019. Moving beyond P values: data analysis  
523 with estimation graphics. *Nature Methods*, 16: 565-566. DOI: [10.1038/s41592-019-0470-3](https://doi.org/10.1038/s41592-019-0470-3).
- 524
- 525 IUCN. 2012. Habitat classification scheme, Version 3.1. *International Union for the Conservation of*  
526 *Nature*. Cambridge, United Kingdom. <https://www.iucnredlist.org/resources/habitat-classification->
- 527 scheme.
- 528
- 529 IUCN. 2021. The IUCN Red List of Threatened Species. Version 2021-2. *International Union for the*  
530 *Conservation of Nature*. Cambridge, United Kingdom. <https://www.iucnredlist.org>.
- 531
- 532 Kassambara A. 2020. ggpubr: 'ggplot2' based publication ready plots. R package version 0.4.0. [online]:  
533 Available from: <https://CRAN.R-project.org/package=ggpubr>.
- 534
- 535 Leclère D, Obersteiner M, Barrett M, Butchart SHM, Chaudhary A, de Palma A, et al. 2020. Bending the  
536 curve of terrestrial biodiversity needs an integrated strategy. *Nature*, 585: 551-556. DOI:  
537 [10.1038/s41586-020-2705-y](https://doi.org/10.1038/s41586-020-2705-y).
- 538
- 539 Ledger S, McRae L, Loh J, Almon R, Böhm M, Clements C, et al. 2021. Past, present and future of the  
540 Living Planet Index; *in preparation*.
- 541
- 542 Leung B, Hargreaves AL, Greenberg DA, McGill B, Dornelas M, and Freeman R. 2020. Clustered versus  
543 catastrophic global vertebrate declines. *Nature*, 267-271. DOI: [10.1038/s41586-020-2920-6](https://doi.org/10.1038/s41586-020-2920-6).
- 544

- 545 Mace GM, Barrett M, Burgess ND, Cornell SE, Freeman R, Grooten M, *et al.* 2018. Aiming higher to bend  
546 the curve of biodiversity loss. *Nature Sustainability*, 1: 448-451. DOI: [10.1038/s41893-018-0130-0](https://doi.org/10.1038/s41893-018-0130-0).
- 547
- 548 Marconi V, McRae L, Müller H, Currie J, Whitmee S, Gadallah F, *et al.* 2021. Population declines among  
549 Canadian vertebrates: But data of different quality show diverging trends. *Ecological Indicators*, 130,  
550 108022. DOI: [10.1016/j.ecolind.2021.108022](https://doi.org/10.1016/j.ecolind.2021.108022).
- 551
- 552 McRae L, Böhm M, Deinet S, Gill M, and Collen B. 2012. The Arctic Species Trend Index: using vertebrate  
553 population trends to monitor the health of a rapidly changing ecosystem. *Biodiversity*, 13:144-156. DOI:  
554 [10.1080/14888386.2012.705085](https://doi.org/10.1080/14888386.2012.705085).
- 555
- 556 Meiri S. 2018. Traits of lizards of the world: Variation around a successful evolutionary design. *Global  
557 Ecology and Biogeography*, 27(10): 1168-1172. DOI: [10.1111/geb.12773](https://doi.org/10.1111/geb.12773)
- 558
- 559 Mooers AØ, Prugh LR, Festa-Bianchet M, and Hutchings JA. 2007. Biases in legal listing under Canadian  
560 Endangered species legislation. *Conservation Biology*, 21(3): 572–575. DOI: [10.1111/j.1523-1739.2007.00689.x](https://doi.org/10.1111/j.1523-<br/>561 1739.2007.00689.x).
- 562
- 563 Myhrvold NP, Baldridge E, Chan B, Sivam D, Freeman DL, and Ernest SKM. 2015. An amniote life history  
564 database to perform comparative analyses with birds, mammals, and reptiles. *Ecology*, 96(11): 3109.  
565 DOI: [10.1890/15-0846R.1](https://doi.org/10.1890/15-0846R.1).
- 566

- 567 Olden JD, Hogan ZS, and Vander Zanden MJ. 2007. Small fish, big fish, red fish, blue fish: size-biased  
568 extinction risk of the world's freshwater and marine fishes. *Global Ecology and Biogeography*, 16(6):  
569 694-701. DOI: [10.1111/j.1466-8238.2007.00337.x](https://doi.org/10.1111/j.1466-8238.2007.00337.x).
- 570
- 571 Oliveira BF, São-Pedro VA, Santos-Barrera G, Penone C, and Costa GC. 2017. AmphiBIO, a global  
572 database for amphibian ecological traits. *Scientific Data*, 4: 170123. DOI: [10.1038/sdata.2017.123](https://doi.org/10.1038/sdata.2017.123).
- 573
- 574 Pata P, Mezzini S, Burant JB, Marconi V, McRae L, Currie J, *et al.* 2020. A sampled approach to the Living  
575 Planet Index; *in preparation*.
- 576
- 577 Pauly D, Christensen V, Guénette S, Pitcher TJ, Sumaila UR, Walters CJ, *et al.* 2002. Towards  
578 sustainability in world fisheries. *Nature*, 418: 689-695. DOI: [10.1038/nature01017](https://doi.org/10.1038/nature01017).
- 579
- 580 RS-eco (2021). traitdata: Easy access to various ecological trait data. R package version 0.0.1. [online]:  
581 Available from: <https://github.com/RS-eco/traitdata>.
- 582
- 583 Saha A, McRae L, Dodd CK, Gadsden H, Hare KM, Lukoschek V, *et al.* 2018. Tracking global population  
584 trends: Population time series data and a Living Planet Index for reptiles. *Journal of Herpetology*, 52(3):  
585 259-268. DOI: [10.1670/17-076](https://doi.org/10.1670/17-076).
- 586
- 587 Santini L, Benítez-López A, Ficetola GF, and Huijbregts MAJ. 2018. Length-mass allometries in  
588 amphibians. *Integrative Zoology*, 13: 36–45. DOI: [10.1111/1749-4877.12268](https://doi.org/10.1111/1749-4877.12268).
- 589

- 590 Sheard C, Neate-Clegg MHC, Alioravainen N, Jones SEI, Vincent C, MacGregor HEA, *et al.* 2020. Ecological  
591 drivers of global gradients in avian dispersal inferred from wing morphology. *Nature Communications*,  
592 11(2463). DOI: [10.1038/s41467-020-16313-6](https://doi.org/10.1038/s41467-020-16313-6).
- 593
- 594 Smith SJ, and Somerton GD. 1981. STRAP: A user-oriented computer analysis system for groundfish  
595 research trawl survey data. *Canadian Technical Report of Fisheries and Aquatic Sciences*, 1030 (4): 1-66.  
596 [online]: Available from: [https://publications.gc.ca/collections/collection\\_2013/mpo-dfo/Fs97-6-1030-eng.pdf](https://publications.gc.ca/collections/collection_2013/mpo-dfo/Fs97-6-1030-eng.pdf).
- 597
- 598
- 599 Stekhoven DJ. 2013. missForest: Nonparametric missing value imputation using random forest. R  
600 package version 1.4.
- 601
- 602 Stekhoven DJ, and Buehlmann P. 2012. MissForest – non-parametric missing value imputation for  
603 mixed-type data. *Bioinformatics*, 28(1): 112-118.
- 604
- 605 Tittensor DP, Walpole M, Hill SLL, Boyce DG, Britten GL, Burgess ND, *et al.* 2014. A mid-term analysis of  
606 progress toward international biodiversity targets. *Science*, 346(6206): 241-244. DOI:  
607 [10.1126/science.1257484](https://doi.org/10.1126/science.1257484).
- 608
- 609 Thorson JT, Cope JM, and Patrick WS. 2014. Assessing the quality of life history information in publicly  
610 available databases. *Ecological Applications*, 24(1): 217-226. DOI: [10.1890/12-1855.1](https://doi.org/10.1890/12-1855.1).
- 611
- 612 Troudet J, Grandcolas P, Blin A, Vignes-Lebbe R, and Legendre F. 2017. Taxonomic bias in biodiversity  
613 data and societal preferences. *Scientific Reports*, 7: 9132. DOI: [10.1038/s41598-017-09084-6](https://doi.org/10.1038/s41598-017-09084-6).

614

615 Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, François R, et al. 2019. Welcome to  
616 the tidyverse. *Journal of Open Source Software*, 4(43): 1686. DOI: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).

617

618 Wilman H, Belmaker J, Simpson J, de la Rosa C, Rivadeneira MM, and Jetz W. 2014. EltonTraits 1.0:  
619 Species-level foraging attributes of the world's birds and mammals. *Ecology*, 95(7): 2027-2027. DOI:  
620 [10.1890/13-1917.1](https://doi.org/10.1890/13-1917.1).

621

622 WWF. 2020. Living Planet Report 2020 - Bending the curve of biodiversity loss. Almond REA. Grooten M.  
623 Petersen T. (Eds). *World Wildlife Fund*. Gland, Switzerland.

624

625 WWF-Canada, ZSL and GFN. 2007. Canadian living planet report 2007. World Wildlife Fund Canada,  
626 Zoological Society of London and Global Footprint Network. Toronto, Canada.

627 [https://www.footprintnetwork.org/content/documents/2007\\_Canadian\\_Living\\_Planet\\_Report.pdf](https://www.footprintnetwork.org/content/documents/2007_Canadian_Living_Planet_Report.pdf).

628

629 WWF-Canada. 2017. Living planet report Canada: A national look at wildlife loss. Currie J, Giles E, and  
630 Snider J. World Wildlife Fund Canada. Toronto, Canada. DOI: [10.13140/RG.2.2.24442.34247](https://doi.org/10.13140/RG.2.2.24442.34247).

631

632 WWF-Canada. 2020. Living Planet Report Canada: Wildlife at risk. Currie J, Snider J, and Giles E. *World*  
633 *Wildlife Fund Canada*. Toronto, Canada. DOI: [10.13140/RG.2.2.16556.49280](https://doi.org/10.13140/RG.2.2.16556.49280).

634

635

636

637

638 **Table Captions**

639 Table 1. Data availability for assessing the representation of species included within the C-LPI, including

640 proportion of vertebrates within Canada, the relative representation of biotic traits.

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664 **Figure Captions**

665 Figure 1. Data availability for assessing the distribution of species traits. Bars depict the representation  
666 of species with trait data in the C-LPI, relative to the number of species with trait data for the broader  
667 vertebrate subphylum in Canada (C-Vertebrates).

668

669 Figure 2. Comparing distributions of (a) body size, (b) lifespan, and (c) trophic level for species included  
670 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
671 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

672

673 Figure 3. Comparing distributions of bird, fish, herpetofauna, and mammal body size for species included  
674 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
675 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

676

677 Figure 4. Comparing distributions of bird, fish, herpetofauna, and mammal lifespan for species included  
678 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
679 (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

680

681 Figure 5. Comparing distributions of bird, fish, herpetofauna, and mammal trophic level for species  
682 included within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking  
683 LPI data (C-Vertebrates Only). Asterisks (\*\*) denote statistical significance.

684

685 Figure 6. Comparing distributions of (a) body size, (b) lifespan, and (c) trophic level for species included  
686 within the C-LPI Trait database, compared to other native Canadian vertebrate species lacking LPI data  
687 (C-Vertebrates Only) using imputed data. Asterisks (\*\*) denote statistical significance.

688

689 **Tables**

690

<b>Data Availability</b>	<b>Birds</b>	<b>Fish</b>	<b>Herps</b>	<b>Mammals</b>	<b>All</b>
C-LPI	388	370	47	101	906
( <i>Representation in C-LPI</i> )	(42.8%)	(40.8%)	(5.2%)	(11.1%)	
C-Vertebrates	453	1044	88	194	1779
( <i>Representation in C-Vertebrates</i> )	(25.5%)	(58.7%)	(4.9%)	(10.9%)	
<b>Species representation (%)</b>	<b>85.7</b>	<b>35.4</b>	<b>53.4</b>	<b>52.1</b>	<b>50.9</b>
C-LPI Trait Database	344	352	46	94	836
( <i>Representation in C-LPI Trait Database</i> )	(41.1%)	(42.1%)	(5.5%)	(11.2%)	
C-Vertebrates Trait Database	404	1007	86	182	1679
( <i>Representation in C-Vertebrates Trait Database</i> )	(24.1%)	(60.0%)	(5.1%)	(10.8%)	
<b>Species trait representation (%)</b>	<b>85.1</b>	<b>35.0</b>	<b>53.5</b>	<b>51.6</b>	<b>49.8</b>
C-Vertebrates Only Trait Database	60	655	40	88	843
( <i>Representation in C-Vertebrates Only Trait Database</i> )	(7.1%)	(77.7%)	(4.7%)	(10.4%)	

691

692

693

694

695

696

697

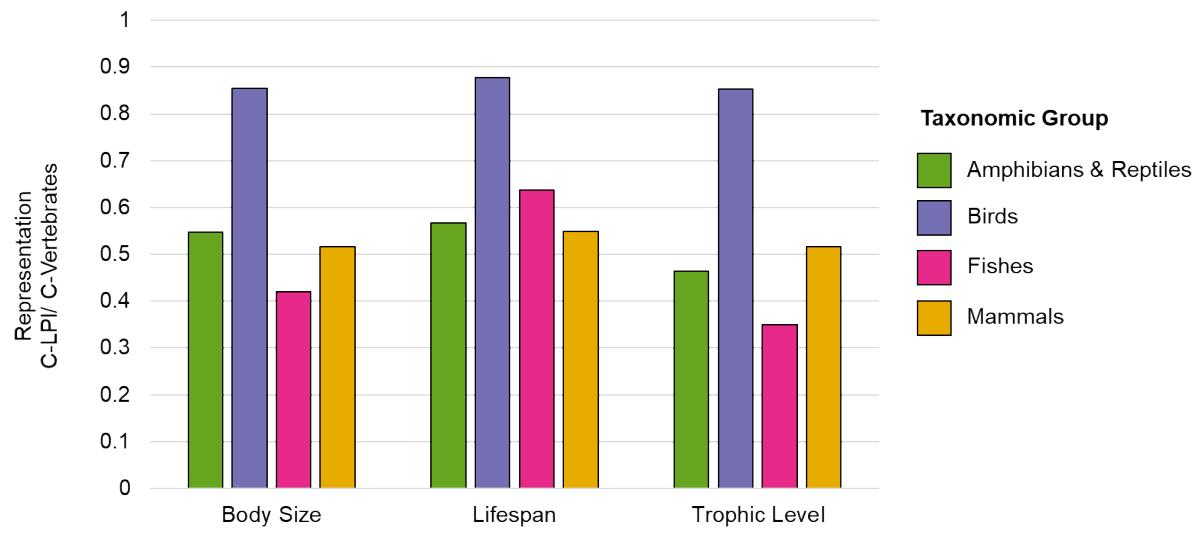
698

699

700

701 **Figures**

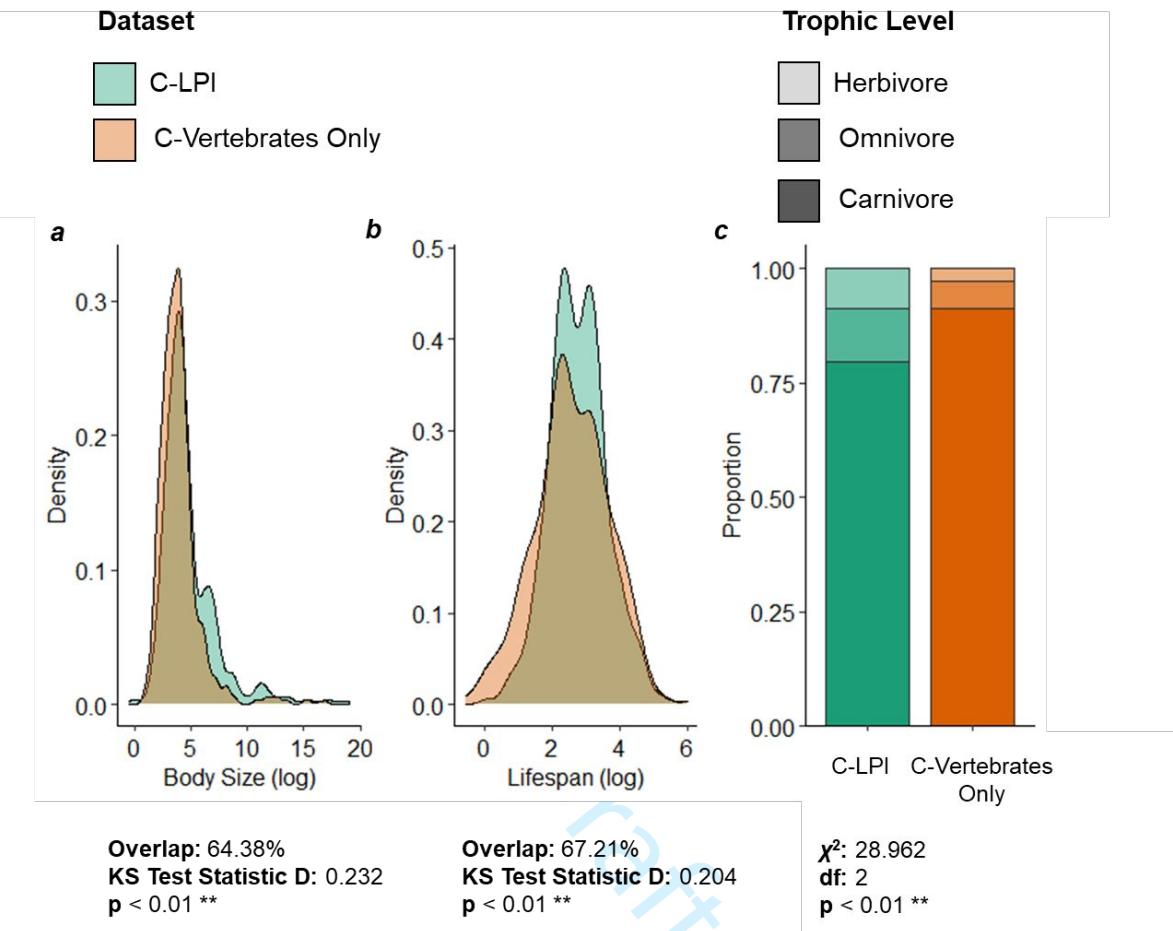
702



703

704 Figure 1.

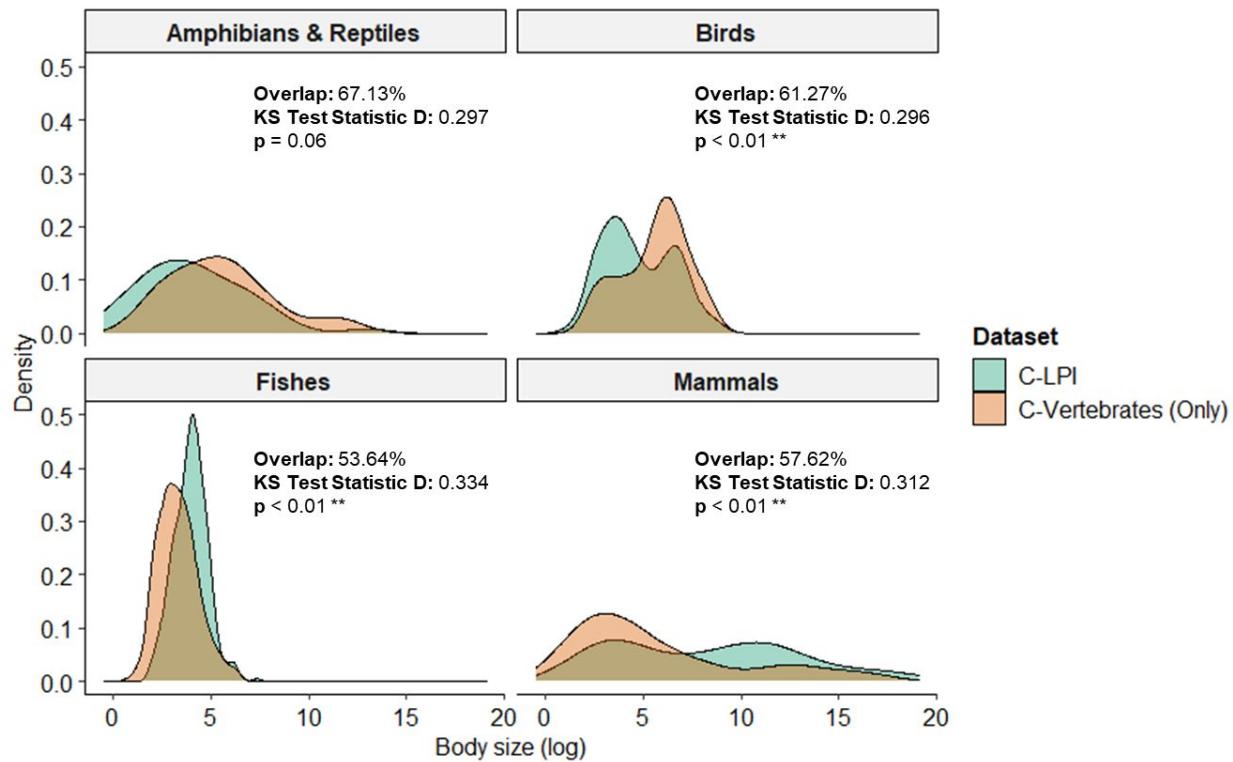
705



706

707 Figure 2.

708

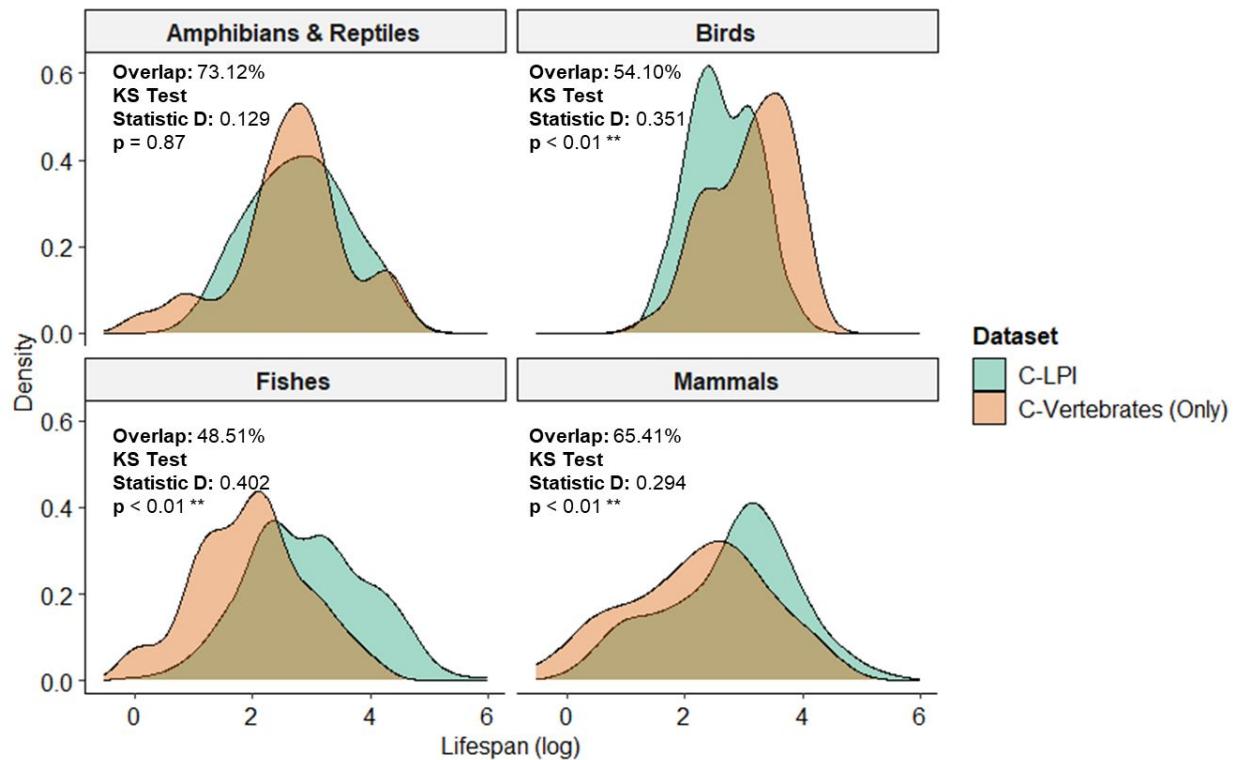


709

710 Figure 3.

711

712

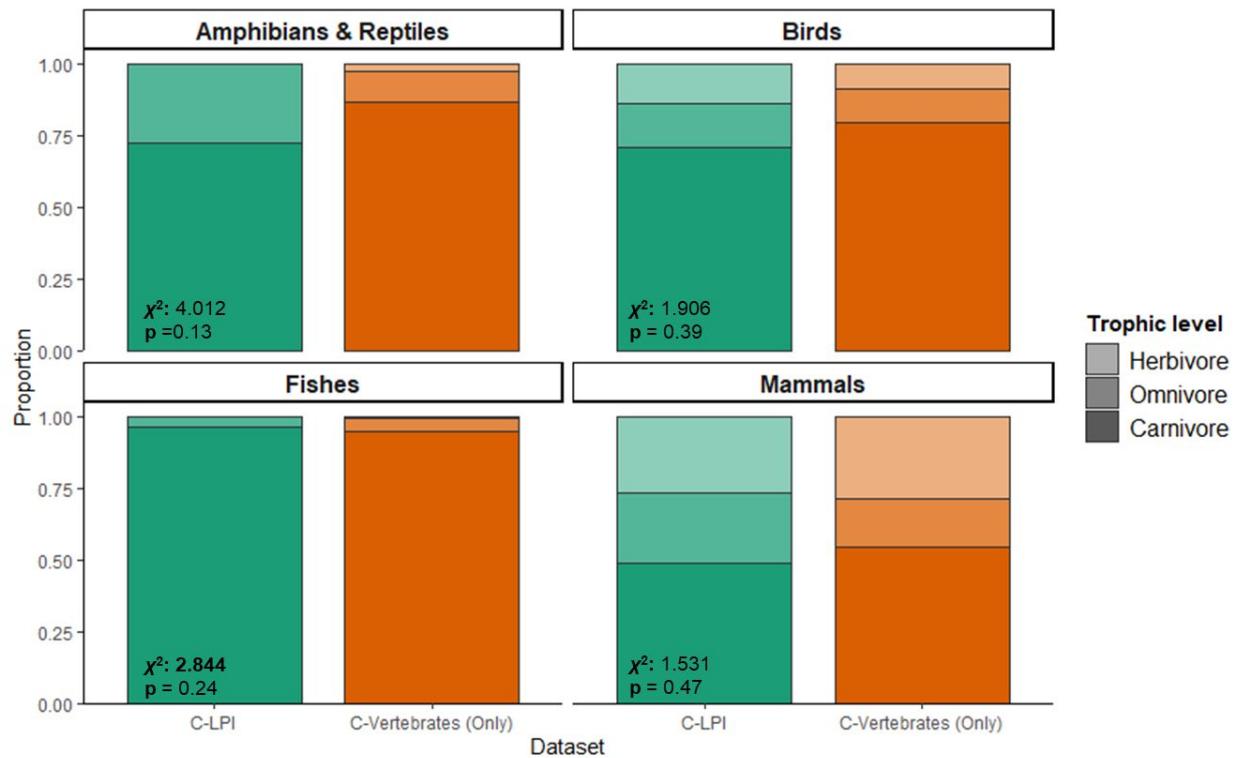


713

714 Figure 4.

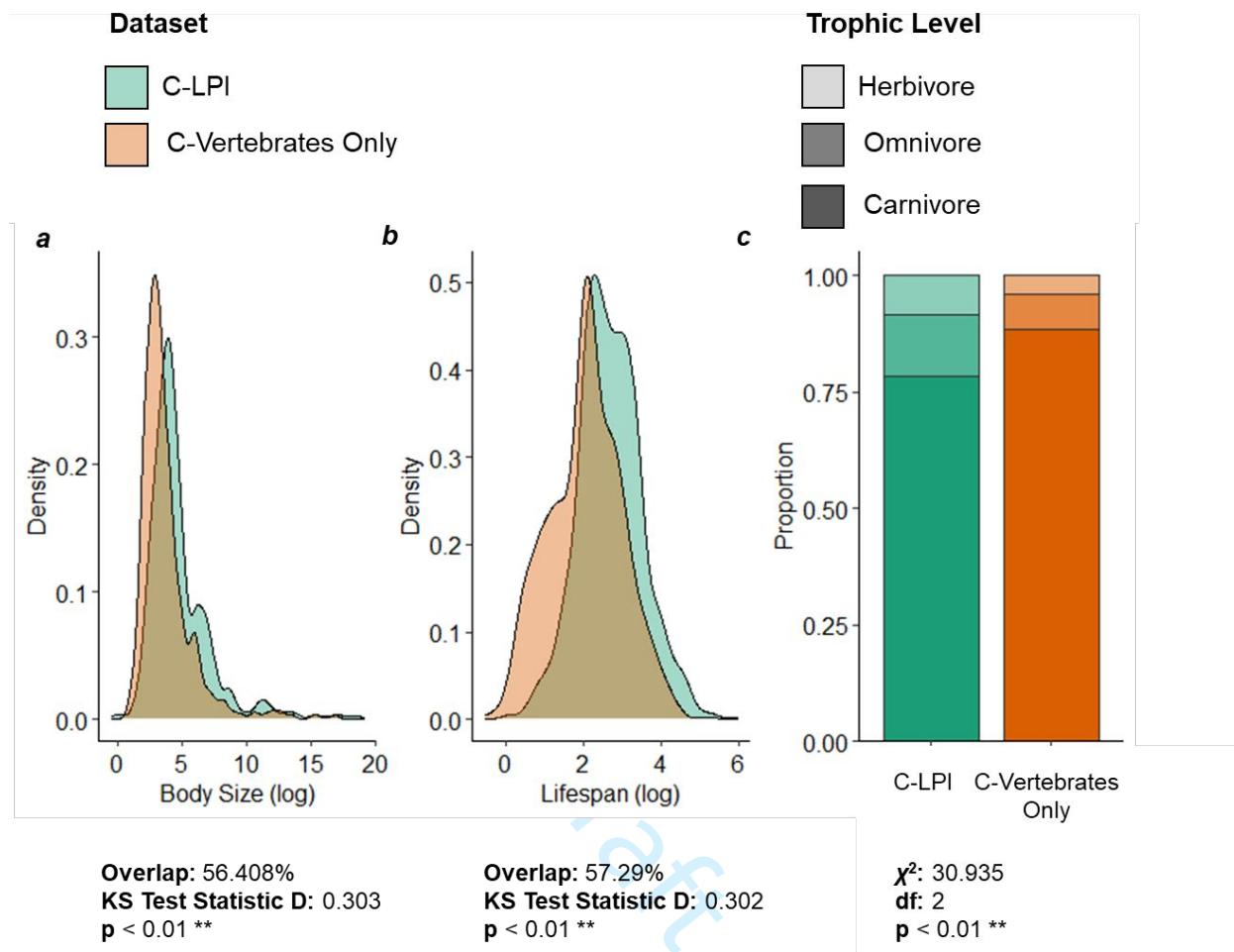
715

716



717

718 Figure 5.



719

720 Figure 6.

<b>Data Availability</b>	<b>Birds</b>	<b>Fish</b>	<b>Herps</b>	<b>Mammals</b>	<b>All</b>
C-LPI <i>(Representation in C-LPI)</i>	388 (42.8%)	370 (40.8%)	47 (5.2%)	101 (11.1%)	906
C-Vertebrates <i>(Representation in C-Vertebrates)</i>	453 (25.5%)	1044 (58.7%)	88 (4.9%)	194 (10.9%)	1779
<b>Species representation (%)</b>	<b>85.7</b>	<b>35.4</b>	<b>53.4</b>	<b>52.1</b>	<b>50.9</b>
C-LPI Trait Database <i>(Representation in C-LPI Trait Database)</i>	344 (41.1%)	352 (42.1%)	46 (5.5%)	94 (11.2%)	836
C-Vertebrates Trait Database <i>(Representation in C-Vertebrates Trait Database)</i>	404 (24.1%)	1007 (60.0%)	86 (5.1%)	182 (10.8%)	1679
<b>Species trait representation (%)</b>	<b>85.1</b>	<b>35.0</b>	<b>53.5</b>	<b>51.6</b>	<b>49.8</b>
C-Vertebrates Only Trait Database <i>(Representation in C-Vertebrates Only Trait Database)</i>	60 (7.1%)	655 (77.7%)	40 (4.7%)	88 (10.4%)	843

Draft

