Université de Montréal

Contributions à l'étude des patrons spatiaux de biodiversité dans les paysages complexes

Par

Julian Wittische

Département de sciences biologiques

Faculté des arts et sciences

Thèse présentée

en vue de l’obtention du grade de *Philosophiae Doctor* (Ph.D.)

en sciences biologiques

June 2022

Université de Montréal

Unité académique : Département de sciences biologiques, Faculté des arts et des sciences

---

Cette thèse intitulée

Contributions à l'étude des patrons spatiaux de biodiversité dans les paysages complexes

Présentée par

Julian Wittische

Sera évaluée par un jury composé des personnes suivantes

Pierre Legendre
Président-rapporteur

Patrick M. A. James
Directeur de recherche

Simon Joly
Membre du jury

Martin-Hugues St-Laurent
Examinateur externe

Résumé

Les patrons spatiaux issus de processus écologiques naturels et anthropiques constituent une mosaïque complexe dans la nature. Ceci combiné aux conséquences parfois colossales des perturbations anthropiques sur la biodiversité et les services écosystémiques, fait de l'écologie spatiale une science à la fois complexe et urgente dans un contexte de changement global rapide. Dans cette thèse, j'ai essayé de contribuer à notre compréhension de comment la structure spatiale des paysages influence la dynamique des populations, en utilisant des approches innovantes. Dans ce but, j'ai étudié différentes perturbations et différents aspects de la dynamique des populations dans trois chapitres. Les deux premiers chapitres d'analyse portent sur un biome qui fournit de nombreux services écosystémiques. La forêt boréale présente une exploitation en plein développement potentiellement non durable, et est aussi actuellement menacée par des perturbations naturelles sans précédents et exacerbées. J'ai d'abord examiné comment l'habitat naturel et les nombreuses altérations du paysage causées par l'homme influencent une communauté de mammifères de la forêt boréale. Ensuite, je me suis concentré sur une autre perturbation à grande échelle de la forêt boréale en identifiant les éléments du paysage qui limitent la propagation d'un des ravageurs forestiers les plus destructeurs au monde. Enfin, j'ai complété les approches utilisées dans les deux chapitres précédents, en me concentrant sur l'aspect temporel du changement de la dynamique des populations. Je l'ai fait en construisant et en évaluant une méthode capable de détecter des changements de diversité génétique locaux et atypiques, malgré les changements aléatoires omniprésents apportés par la dérive génétique et le flux de gènes. Mes trois chapitres d'analyse ont des implications claires en matière de conservation. Bien que deux d'entre eux soient concentrés sur des systèmes spécifiques, ils peuvent s'appliquer à d'autres paysages, ou du moins fournir une piste pour de futures recherches. Pour conclure ma thèse, j'ai suivi la synthèse de mes chapitres par une discussion sur comment l'interdisciplinarité des disciplines associées à l'écologie spatiale est une force essentielle que nous devrions cultiver. Je termine ma thèse en identifiant certaines des directions de recherche futures les plus excitantes et prometteuses.

Mots-clés : Écologie spatiale ; Structure du paysage ; Perturbation ; Modélisation des communautés ; Génétique des populations ; Forêt boréale ; Épidémies d’insectes ; Changement temporel ; Biodiversité

Abstract

Natural and anthropogenic spatial patterns create an intricate mosaic. This fact, combined with the sometimes colossal consequences of anthropogenic disturbances on biodiversity and ecosystem services, makes spatial ecology a science of both complexity and urgency in a context of rapid global change. In this thesis, I have tried to contribute to our understanding of how the spatial structure of landscapes influences population dynamics through innovative approaches. Towards this goal, I have investigated different perturbations and different aspects of population dynamics in three chapters. The first two analysis chapters focus on a biome which is a large provider of ecosystem services and resources, featuring rapidly increasing and possibly unsustainable exploitation, but which is also currently under threat from unprecedented and exacerbated natural disturbances. First, I delved into how a community of boreal forest mammals is driven by its natural habitat and the many man-made alterations to the landscape. Then, I focused on another large-scale perturbation of the boreal forest by identifying what elements of the landscape constrain the spread of one of the most destructive forest pests in the world. Finally, I complemented the approaches used the previous two chapters, by focusing on the temporal aspect of population dynamics change. I did this by constructing and evaluating of a method capable of detecting local atypical change in genetic diversity despite the ever-present random changes brought by genetic drift and gene flow. My three analysis chapters have clear conservation implications which, although two of them focused on specific systems, may translate to other landscapes, or at least provide a pipeline for future research. To conclude my thesis, I followed the synthesis of my chapters by a discussion about how the interdisciplinarity of disciplines associated with spatial ecology is an essential strength we should cultivate. I end my thesis by identifying some of the most exciting and promising future research directions.

Keywords: Spatial ecology; Landscape structure; Disturbance; Community modelling; Population genetics; Boreal forest; Insect outbreaks; Temporal change; Biodiversity

Table des matières

Liste des tableaux

Liste des figures

Liste des matériels supplémentaires

Liste des abréviations

3D	*Three-dimensional* Tridimensionnel
AB	Alberta
ABMI	*Alberta biodiversity monitoring institute* Institut de surveillance de la biodiversité d'Alberta
ADN / DNA	Acide désoxyribonucléique / *Deoxyribonucleic acid*
AIC	*Akaike information criterion* Critère d'information d'Akaike
AVI	*Alberta vegetation inventory* Inventaire de la végétation d'Alberta
BC	*British Columbia* Colombie-Britannique
CFS	*Canadian forest service* Service canadien des forêts
CI	*Confidence interval* Intervalle de confiance
CRSNG/NSERC	Conseil de recherches en sciences naturelles et en génie du Canada / *Natural Sciences and Engineering Research Council of Canada*
CSI	*Climate suitability index* Index de convenance climatique
CTA	*Community trajectory analysis* Analyse de trajectoire des communautés
DPP / MPB	Dendroctone du pin ponderosa / *Mountain pine beetle*

FNR	*False negative rate* Taux de faux négatifs
FPR	*False positive rate* Taux de faux positifs
F_{ST}	*Fixation index* Indice de fixation
GA	*Genetic Algorithm* Algorithme génétique
GIS	*Geographic information system* *Système d'information géographique*
GMTED	*Global multi-resolution terrain elevation data* Données globales multi-résolutions d'altitude
GWAS	*Genome-wide association studies* Étude d'association pangénomique
IPCC	*Intergovernmental Panel on Climate Change* Groupe d'experts intergouvernemental sur l'évolution du climat
LCBD	*Local contributions to beta diversity* Contributions locales à la diversité bêta
LL	*Log-likelihood* Log-vraisemblance
MCMC	Markov chain Monte Carlo Méthode de Monte-Carlo par chaînes de Markov
MEM	*Moran's eigenvector map* Carte de vecteurs propres de Moran
MLPE	*Maximum-likelihood population-effects (models)* (Modèles) basés sur le maximum de vraisemblance et prenant en compte les effets de population

NRCAN	*Natural resources Canada* Ressources naturelles Canada
R^2	*Coefficient of determination* Coefficient de détermination
RDA	*Redundancy analysis* Analyse de redondance
SNP	*Single-nucleotide polymorphism* Polymorphisme d'un seul nucléotide
SRTM	*Shuttle radar topography mission* Mission topographique par inférométrie radar sur navette
TBI	*Temporal beta-diversity indices* Indices de diversité bêta temporels
TGI	*Temporal genetic diversity indices* Indices de diversité génétique temporels
UALF	*University of Alberta Linear Features* University of Alberta – Éléments linéaires

Remerciements

Cette thèse n'a été possible qu'avec le concours de nombreuses personnes. J'ai eu la chance de rencontrer un grand nombre de personnes gentilles et qualifiées qui m'ont permis de construire ma thèse et de la soutenir aujourd'hui. Il est donc de mon devoir de remercier tous ceux qui ont participé pour leurs leçons, leur aide, leur support, et leur encouragement.

Il s'agit en particulier de mon superviseur, Patrick James, qui a fait preuve d'une patience et d'une flexibilité miséricordieuses à mon égard. Il est indéniable qu'il continuera à influencer mon futur dans la recherche. Il y a également d'autres chercheurs, trop nombreux pour tous les nommer, qui ont contribués à mon développement en tant que scientifique : Jeremy Larroque, Ronan Marrec, Louis-Étienne Robert, Pierre Legendre, Simon Joly, Jesse Shapiro, Steven Kembel, Daniel Borcard, Bernard Angers, Simon Gravel, Jasmine Janes, Colin Garroway, Scott Heckbert, Jason Fisher, Cole Burton… Ceci est sans compter les chercheurs qui m'avaient stimulé avant mon doctorat : David Wake, Isabelle Olivieri, Jonathan Levine, Fabien Aubret, Fabien Leprieur, Michelle Koo, Sabine, Güsewell, Carole Spencer, Daniel Rabosky …

Il s'agit bien entendu aussi d'autres étudiants-chercheurs, trop nombreux pour tous les nommer, avec qui j'ai eu la chance d'avoir des échanges scientifiques fructueux : Simon Legault, Alexis Carteron, Paul Mayrand, Olivier Pontbriand-Paré, Érik L'Heureux, Vincent Chapdelaine-Trépanier, Maëva Perez, Patricia Sujii, Jennifer Sauri Ramirez, Daiana Tourne, Ryan Ju, Chloé Frédette, Salomé Bouskila, Mathilde Léonard …

Il ne faut surtout pas oublier tous les amis qui m'ont permis de surmonter les périodes sombres de ma thèse : en premier lieu Hinatea Ariey, mais aussi Charlotte Van Engeland, Julie Marleau, Flavian Mèche, Caitlin McGee, Claire Barelli, Clémence Ragon, Noémie Le Jeune, Rayän El Ghoneimi, Rosie Cochrane, Maïté Rohaut, John Nguyen …

Merci également à tous les organismes qui ont contribués au financement de mes études : CRSNG (TRIANet, FONCER), l'Université de Montréal, le CSBQ, la CSEE et le CEF.

Finalement, je voudrais remercier mes parents de m'avoir mis au monde, et de m'avoir réanimé 7 jours après ma naissance pendant une mort subite du nourrisson, et pour leur amour inconditionnel et leur irréductible soutien.

À propos du raisonnement et de la résolution de problèmes en sciences :

« *[…] quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo (accito si placet amico) dicere: Calculemus*. »

« *[…]si des controverses surgiraient, il n'y aurait pas plus besoin de plus discussion entre deux philosophes qu'entre deux mathématiciens. Car il leur suffirait de prendre leurs plumes dans leurs mains et de s'asseoir devant un abaque[1], et de se dire entre eux (et pourquoi pas aussi à un ami qu'ils inviteraient à l'aide) : Calculons !* »

« *[...] if controversies were to arise, there would be no more need of disputation between two philosophers than between two calculators. For it would suffice for them to take their pencils in their hands and to sit down at the abacus[1], and say to each other (and if they so wish also to a friend called to help): Let us calculate* »

Gottfried Wilhelm Leibniz

[1] un outil qui sert à calculer

Un petit poème que m'a inspiré cette thèse :

Très chers collègues écologues,
Pas de répit pour notre récit,
Car nos complexes épilogues,
Sont souvent d'humbles prologues,
Innés ennemis de l'ennui !

1. CHAPITRE 1 – INTRODUCTION GÉNÉRALE

1.1 L’écologie : une science de la complexité

La discipline qu’est l’écologie commença à prendre forme au début du 20e siècle et connut un essor ininterrompu jusqu’ici. La description de l’histoire naturelle, de la biodiversité et des phénomènes écosystémiques, vit des contributions majeures dès le 18e siècle par Buffon, de Candolle et Humboldt, parmi d’autres (Acot et al. 1998; Ebach & Goujet 2006; Nelson 1978). La théorie de la sélection naturelle (Darwin & Wallace 1858), en répondant à certaines questions concernant les origines de la biodiversité, stimula encore plus la recherche concernant la coexistence des espèces et leur répartition. L’écologie fut définie à l’origine par Ernst Haeckel comme étant la science de la relation des organismes avec leur environnement biotique et abiotique (Egerton 2013; Haeckel 1866), et se base donc sur cette histoire naturelle en mettant l’emphase sur la quantification et la compréhension des patrons et processus à l’origine de cette biodiversité. La nécessité du maintien de la biodiversité pour le fonctionnement des écosystèmes devint plus claire à mesure que l’écologie connût un raffinement dans la compréhension des patrons de biodiversité et de leurs changements. L’avènement de l’actuel changement global a rendu ces connaissances essentielles.

L’écologie est une science de la complexité, celle du fonctionnement des écosystèmes, celle de la dynamique des populations et des communautés. Parmi les nombreux concepts apparus au début du développement de l’écologie figurent la niche écologique (Elton 1927; Gravel et al. 2019; Grinnell 1917; Hutchinson 1957), l’exclusion compétitive (Gause 1934), la théorie de la biogéographie insulaire (MacArthur & Wilson 1967), la théorie du chaos (May 1974), et la métapopulation (Hanski 1994). Les progrès dans notre compréhension de l’évolution ont également permis de mieux conceptualiser les relations entre les espèces, notamment grâce au concept de coévolution (Ehrlich & Raven 1964). Cela dit, l’utilisation de termes et concepts en écologie et dans les sciences associées ne fut pas exempte de confusion (Johnson & Lidström 2018; Tansley 1935). Encouragée par cette richesse conceptuelle, s'en suivit l'application de plus en plus fréquente de la théorie de manière quantitative, sur les mesures issues de l'observation de terrain ou expérimentale.

La complexité en écologie est également celle des approches méthodologiques. Par exemple, la description statistique des tests d'hypothèses nulles, à travers les *p*-valeurs, connut un grand essor durant le siècle dernier (Low-Décarie et al. 2014), jusqu'à devenir très fréquent dans les publications. La meilleure approche à utiliser pour analyser et interpréter les données écologiques fait toujours débat, par exemple en ce qui concerne la signification statistique et la reproductibilité (Ellison 2004; Halsey et al. 2015; Verhulst 2014; White et al. 2014). Une autre avenue méthodologique qui a connu un développement très riche et qui a permis d'utiliser des données de plus en plus complexes est celle de l'analyse multivariée (Legendre & Legendre 2012), qui se développa rapidement à partir des années 50, notamment pour répondre à des questions concernant l'écologie des communautés (Odum 1950) et la classification (Goodall 1954). Depuis, l'analyse numérique s'est grandement diversifiée parallèlement au raffinement et à la diversification des questions, et à la nécessité d'inférences plus robustes (Legendre & Legendre 2012). Le dynamisme de la pratique de l'inférence en écologie est toujours d'actualité avec de nombreux et fructueux débats académiques.

1.2 La complexité en écologie : entre muse et fardeau

L'écologie devient de plus en plus essentielle pour l'humanité dans un monde qui se complexifie et change de plus en plus vite. Nous sommes actuellement en pleine crise biotique : la sixième grande extinction (Barnosky et al. 2011; Ceballos et al. 2015; Dirzo et al. 2014; Humphreys et al. 2019; Vellend et al. 2017). De nombreuses espèces à travers tous les règnes du vivant se sont récemment éteintes ou ont connu des déclins importants à cause de l'action directe ou indirecte de l'Homme. Parmi les actions directes (Ripple et al. 2019) se trouvent entre autres la surchasse (Benítez-López et al. 2017, 2019; Darimont et al. 2015; Redford 1992), la surpêche (Dulvy et al.

2003; Hilborn 2012; McCauley et al. 2015), le braconnage (Gavin et al. 2010; Liberg et al. 2012; Poulsen et al. 2017), et la destruction des habitats (Foley et al. 2005; Maxwell et al. 2016; Pimm et al. 2014). Les actions humaines indirectes, mais tout aussi néfastes pour la biodiversité sont, entre autres, la fragmentation des habitats (Fahrig 2003a; Gibson et al. 2013), la pollution (Dise et al. 2011; Hölker et al. 2010; Wood et al. 2010), les introductions et invasions y compris des maladies (Bellard et al. 2016; Doherty et al. 2015; Fisher & Garner 2020) et le changement climatique (Bellard et al. 2012) qui éteint déjà des espèces indépendamment des autres facteurs (Waller et al. 2017). Nous ne sommes pas encore capables de contrer ces déclins efficacement et à temps (Butchart et al. 2010; Rands et al. 2010; Rockström et al. 2009). Malheureusement, l'activité et la consommation humaine future vont sans doute continuer à contribuer au déclin des espèces et écosystèmes (Bellard et al. 2012; Pereira et al. 2010; Powers & Jetz 2019; Ripple et al. 2019). Au-delà de stimuler la recherche scientifique, ce contexte de changements biotiques et abiotiques importants rend l'écologie nécessaire pour initier et guider des actions de conservations efficaces (Phillis et al. 2013; Sinclair et al. 2018; Szabó & Hédl 2011).

Grâce à la découverte de nouveaux niveaux de biodiversité et d'interactions entre individus et espèces, les écosystèmes ont été reconnus comme des systèmes très complexes. Cette complexité peut être théorisée à travers les concepts d'auto-organisation et d'imprévisibilité, ce qui peut aider les gestionnaires à mieux guider leurs actions (Newman et al. 2019; Parrott 2002). La reconnaissance de cette complexité a aussi changé le raisonnement scientifique en écologie en promouvant la synthèse, et en maintenant les échanges, entre les différentes branches de l'écologie (Naeem 2002). En effet, la complexité peut créer une divergence dans l'utilisation de concepts et ressources entre sous-disciplines de l'écologie (DiLeo & Wagner 2016; Storfer et al. 2007). Finalement, identifier de nouvelles sources d'incertitude, qui est prépondérante et variée en écologie, permet de mieux raffiner les concepts (Milner-Gulland & Shea 2017; Yanai et al. 2018), repoussant ainsi les frontières scientifiques en écologie.

Les frontières scientifiques en écologie sont certes en constante expansion, grâce aux nouveaux mécanismes découverts et à la disponibilité de nouvelles variables et méthodes, mais il reste des lacunes à combler. Les avances conceptuelles et technologiques ne signifient pas nécessairement que le pourcentage de variance des phénomènes naturels expliqué augmente de manière linéaire avec le temps (Low-Décarie et al. 2014). L'écologie présente également un apparent conflit entre la simplification inhérente aux concepts théoriques, et l'important degré d'incertitude et d'erreur présent dans les phénomènes écologiques (Johnson & Lidström 2018). Le risque est de trop simplifier ou homogénéiser l'interprétation écologique, ce qui risque une mauvaise communication des conclusions aux gestionnaires de ressources naturelles (Gallagher & Appenzeller 1999; Harwood & Stokes 2003; Lakoff 2010). Une des additions conceptuelles les plus considérables pour mieux comprendre les changements de biodiversité et améliorer leur interprétation fut l'intégration de l'espace géographique et du concept de paysage.

1.3 Le complexe paysage scientifique du paysage en écologie

La première loi de la géographie (Tobler 1970) indique que « Tout est lié à tout le reste, mais les choses proches sont plus liées que les choses éloignées. », et cette notion est au cœur de la modélisation de la plupart des processus spatiaux en écologie. Les forces endogènes et exogènes qui peuvent générer des patrons spatiaux sont très nombreuses et variées, et existent à toutes les échelles (**Figure 1-1**). En effet, les processus inhérents aux espèces tels que le flux de gènes étudié dans le Chapitre 3, par exemple, peuvent générer une corrélation entre mesures faites sur des individus proches géographiquement—ou autocorrélation spatiale. L'environnement peut également induire une dépendance spatiale par l'hétérogénéité spatiale des variables influant, par exemple, sur l'activité des individus telle qu'étudiée dans le Chapitre 2. Cette prise de conscience conceptuelle a été très importante pour la modélisation en écologie. En effet, les modèles prenant en compte la position des observations dans le paysage, et donc la dépendance et l'autocorrélation spatiales, apparurent plus pertinents (Bascompte & Sole 1996; Duning 1995; Turner et al. 1995). L'échelle temporelle se superpose à l'échelle spatiale et fait partie intégrale de nombreuses questions en écologie du paysage, comme celle explorée dans le Chapitre 4, car elle est à la base du plan d'échantillonnage (Preston 1960), de la notion de changement intrinsèque des populations (Bradburd & Ralph 2019; Fenderson et al. 2019), des cycles de perturbation (James et al. 2011b), et des communautés , et de la relation entre les hommes et leur environnement (Balée & Erickson 2016).

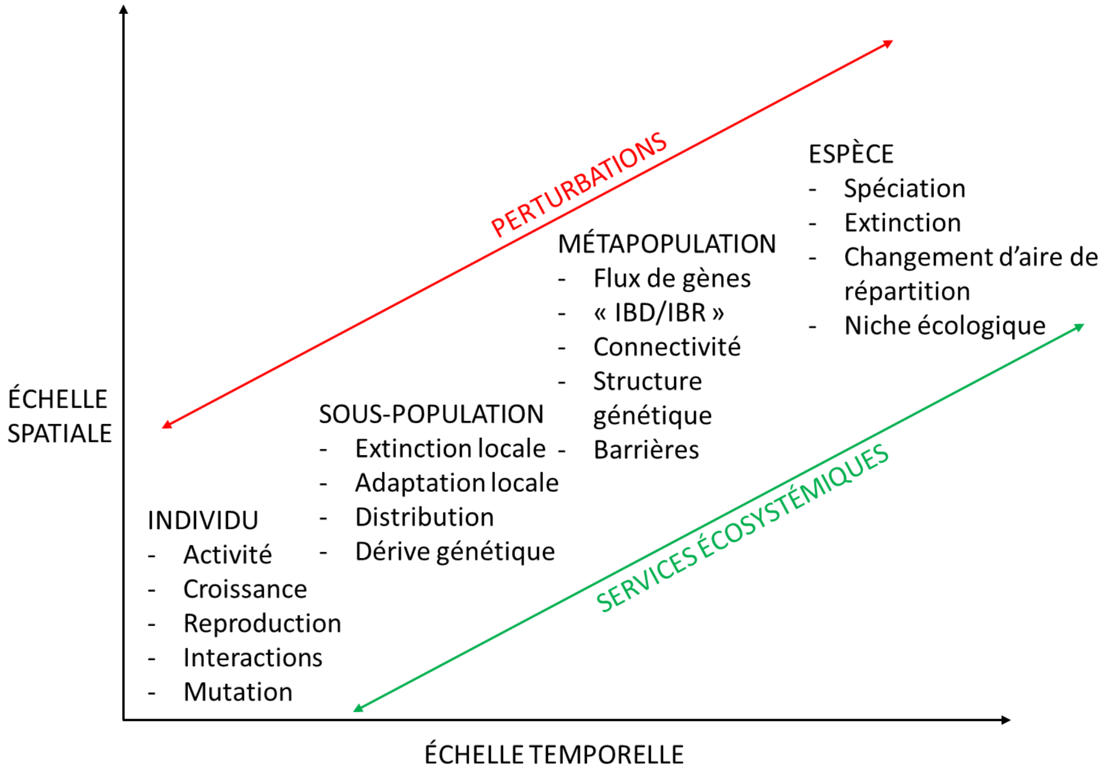


Figure 1-1 : Exemples de processus et de patrons associés à différentes échelles spatiales et temporelles.

Le concept de paysage est complexe, mais central à de nombreuses questions en écologie spatiale (Turner 1989, 2005, Wiens 1992, Manel et al. 2003; **Figure 1-2**), et fournit un concept unificateur aux questions abordées dans cette thèse. Le paysage utilisé dans une étude, pour répondre à une question spatiale ou spatio-temporelle, devrait être défini a priori à partir d'hypothèses spécifiques à ce paysage (Newman et al. 2019). Définir un paysage implique généralement une liste de variables, une proposition de la relation fonctionnelle—une fonction linéaire est (trop ?) souvent considérée—entre ces variables et le phénomène étudié dans une étendue délimitée et à un grain (ou résolution) spécifique (Turner 2005). Les organismes sont affectés par les propriétés de leur paysage, qui est spatialement structuré à de nombreuses échelles (Levin 1992) et qui est spatialement et temporellement variable (Kareiva & Wennergren 1995; Pickett & Cadenasso 1995). Il est difficile de distinguer clairement les liens et rétroactions entre processus et patrons écologiques derrière ces propriétés, car la relation entre processus et patron n'est pas nécessairement monotone dans le paysage, car différents processus peuvent s'influencer entre eux, ou créer des patrons similaires (Dale & Fortin 2014; Turner 1989).

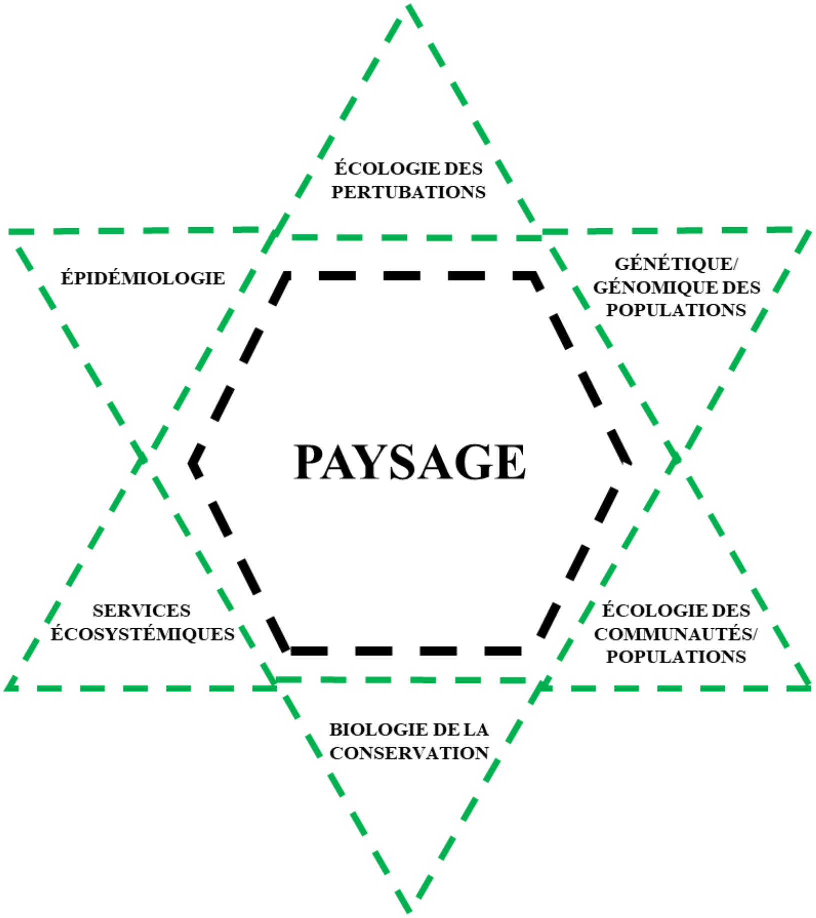


Figure 1-2 : Quelques disciplines liées à l'écologie pour lesquelles le concept de paysage peut être central.

Cette relation complexe entre processus et patron a été spécialement mise en avant par l'écologie des perturbations, qui a fortement influencé le développement de l'écologie du paysage (Turner 2005). La localisation des organismes et de leurs groupements (**Figure 1-1**), ainsi que la distance absolue ou relative entre eux, sont des données clés pour comprendre les interactions écologiques au niveau du paysage (Turner 1989; Wiens 1992). La connaissance des processus à l'origine des patrons écologiques peut permettre de mieux choisir l'échelle à laquelle étudier ces patrons, ou du moins à laquelle ils sont interprétables (Dale & Fortin 2014; Estes et al. 2018; Mayor et al. 2009; Wiens 1989). Il est possible de définir trois grandes catégories de questions qui s'intéressent à l'effet du paysage sur la dynamique des populations (DiLeo & Wagner 2016; Gardner & Urban 2007; Wagner & Fortin 2013; With et al. 1997), qui représente l'axe de recherche central de cette thèse. Tout comme pour d'autres catégorisations de questions en écologie du paysage, ces questions ne sont pas mutuellement exclusives. Les réponses à ces questions se complètent souvent pour avancer vers un but de conservation, mais demandent chacune des approches particulières.

La première grande catégorie de questions inclut les questions qui se focalisent sur l'effet de la qualité et la quantité de l'habitat local sur diverses caractéristiques locales des sous-populations (**Figure 1-3**.-A). On parle alors d'analyse basée sur les nœuds (Dale & Fortin 2010; Fall et al. 2007; Wagner & Fortin 2013). Les données qui servent de réponses dans les modèles utilisés pour répondre à ces questions peuvent par exemple inclure des mesures de diversité (p. ex., génétique, spécifique…), des mesures d'abondance, ou encore des mesures d'activité. Les données qui servent de prédicteurs sont mesurées sur les sites d'échantillonnages, mais elles peuvent également inclure une mesure qui résume les caractéristiques d'une zone plus ou moins grande autour de la sous-population focale (on peut alors aussi parler d'analyse basée sur le voisinage). Ces questions sont très variées et peuvent inclure des questions portant sur la distribution, un résumé focal de la connectivité locale, ou encore l'adaptation locale. Le Chapitre 2 fait partie de cette catégorie de questions, car il se concentre sur l'évaluation de l'effet de la quantité et la qualité de l'habitat local sur des sous-populations. Le Chapitre 2 utilise de nombreuses sources d'hétérogénéité biotique et de perturbation anthropique abiotique pour expliquer l'activité locale de nombreux animaux.

La deuxième grande catégorie de questions inclut les questions qui se focalisent sur l'effet de l'habitat sur les échanges entre les sous-populations (**Figure 1-3**.-B). On parle alors d'analyse basée sur les liens. Les données qui servent de réponses dans les modèles utilisés pour ces questions sont donc généralement associées à une connectivité fonctionnelle (Taylor et al. 1993) entre les sous-populations. Il s'agit donc souvent de variables associées au mouvement et/ou au flux de gènes entre les sous-populations qui sont calculées comme dissimilarités ou distances entre toutes les paires de sous-populations du paysage. Ces mesures, par exemple des distances écologiques, ne peuvent donc pas être indépendantes les unes des autres pour un même paysage (Clarke et al. 2002). La matrice du paysage peut affecter les échanges entre sous-populations de maintes façons (Bonte et al. 2012), et il est possible de développer des hypothèses à propos de comment la connectivité structurelle influence le mouvement dans le paysage à partir de la composition et de la configuration de la matrice (Taylor et al. 1993). La construction de ce type de modèle doit donc prendre en compte le comportement de dispersion des individus dans le cas des animaux à dispersion active (Baguette et al. 2013; Clobert et al. 2009). Certaines études parviennent à combiner ces deux approches, au niveau des nœuds et entre les nœuds, en utilisant une approche de modèles de gravité (Dileo et al. 2014; Murphy et al. 2010) ou de modèles de fonction d'incidence (Graham et al. 2018). Le Chapitre 3 fait partie de cette catégorie de questions, car nous évaluons l'effet de l'environnement, tel que le climat et l'altitude par exemple, sur les échanges génétiques entre des sous-populations d'un animal suivant une explosion démographique.

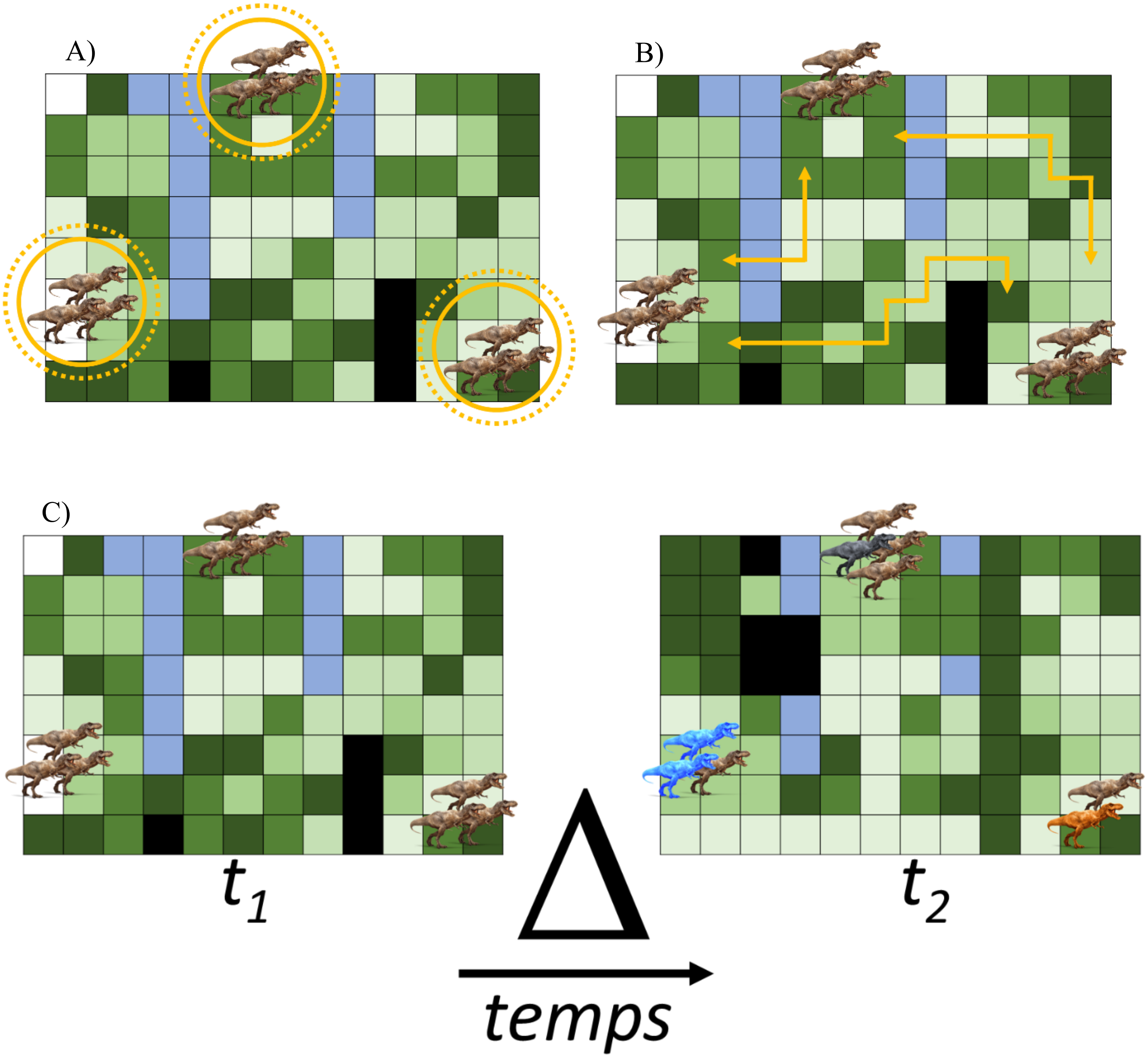


Figure 1-3 : Différentes approches pour quantifier l’effet du paysage sur des populations. A) Ce type de question se concentre sur l’effet local du paysage sur les sous-populations (nœuds), le Chapitre 2 en est un exemple ; B) Ce type de question se concentre sur l’effet du paysage entre les nœuds (liens) sur les mouvements entre ceux-ci, le Chapitre 3 en est un exemple ; C) Ce type de question se concentre sur les changements temporels dans le paysage, le Chapitre 4 en est un exemple.

La troisième grande catégorie de questions centrées sur le paysage est celle qui met l'accent sur la dimension temporelle des problématiques écologiques (**Figure 1-3**.-C). Il s'agit souvent de décrire et comprendre le changement temporel dans les propriétés des populations et/ou du paysage. La quantification précise des changements temporels de diversité, génétique ou spécifique, présente néanmoins de nombreux défis (Borcard et al. 2018; Dornelas et al. 2013). Par exemple, il est complexe d'étudier l'effet d'un processus écologique particulier sur la démographie d'une espèce, car de multiples changements se produisent en permanence à une multitude d'échelles temporelles (Ontiveros et al. 2021), et il faut s'assurer que l'échelle temporelle du processus démographique étudié soit cohérente avec l'étendue temporelle de l'étude (Anderson et al. 2010; Estes et al. 2018). Décrire les trajectoires des communautés est un outil particulièrement intéressant pour résumer ces changements (De Cáceres et al. 2019; Sturbois et al. 2021). La propriété transgénérationnelle des données génétiques les rend très utiles pour étudier les changements temporels. Par exemple, prendre en compte les processus coalescents en génétique du paysage peut permettre de mieux cerner les changements démographiques dans le paysage (Leblois et al. 2014), ce qui permet de construire des pedigrees spatiaux qui peuvent décrire de manière très précise les mouvements à travers les générations (Bradburd & Ralph 2019). Évaluer la relation entre le paysage et la structure génétique au cours du temps permet de suivre la réponse des populations aux perturbations (Draheim et al. 2018; Holzhauer et al. 2006), et éventuellement de prédire les changements démographiques futurs (Fenderson et al. 2019). Cela permet également de discriminer effets contemporains et historiques du paysage sur les propriétés de la population (Pavlacky et al. 2009). Un problème récurrent, peu importe la catégorie de questions auxquelles on s'intéresse, est de savoir comment obtenir les données qui serviront de réponses. En effet, il est crucial d'observer ces phénomènes naturels, qui peuvent se révéler importants pour la gestion des écosystèmes et des espèces menacées, mais qui sont souvent difficiles à suivre par observation directe.

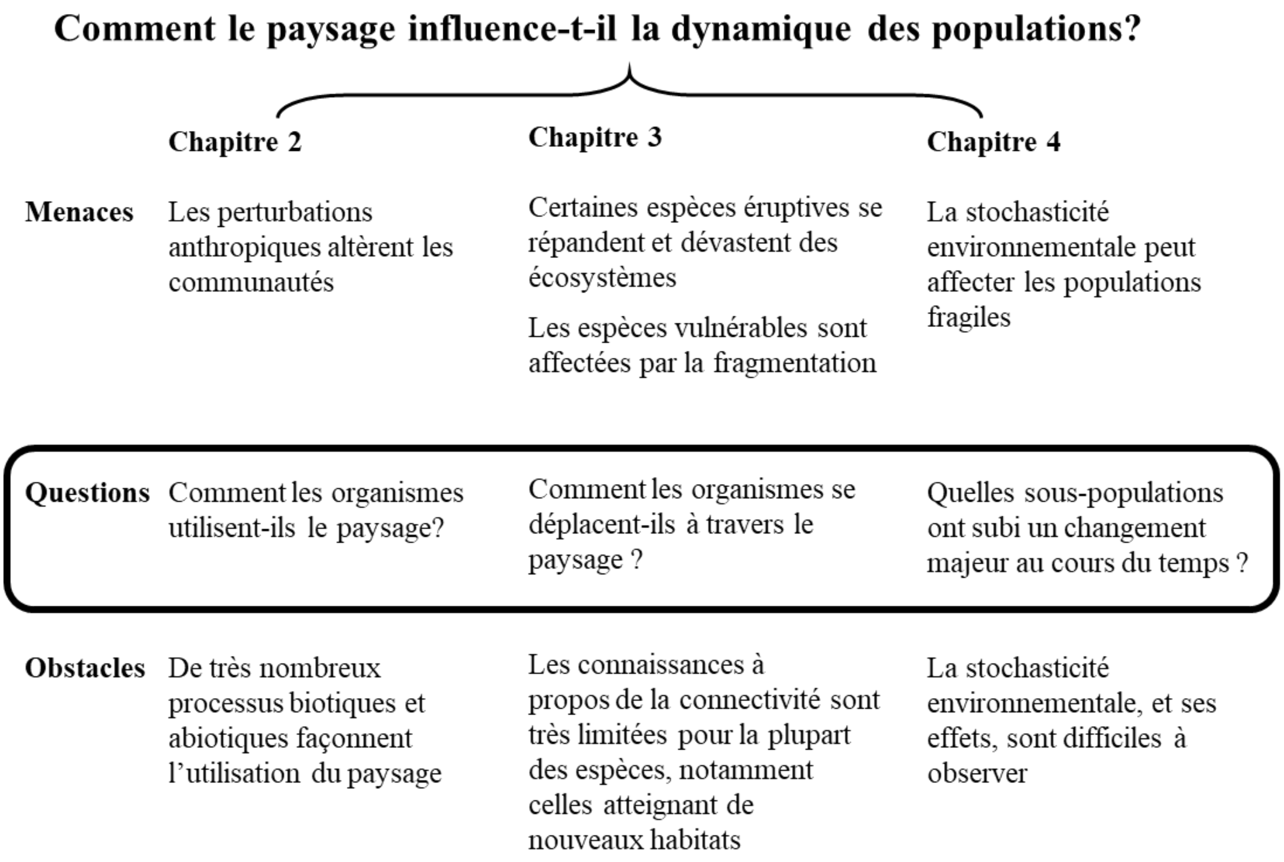
1.4 Quantifier l’inobservable

Les mesures indirectes des propriétés des sous-populations sont devenues un outil inestimable dans la pratique moderne de l’écologie. En effet, dans le cas des animaux, il est très difficile d’observer de nombreux individus sur une longue durée, et il n’est pas toujours possible technologiquement ou financièrement de les suivre à tout instant (Hastings & Harrison 1994; Ims & Yoccoz 1997). C’est notamment le cas pour de nombreux invertébrés (Osborne et al. 2002) et les espèces rares, menacées (Legge et al. 2018) ou inaccessibles (Thompson 2004). Parmi les méthodes qui permettent d’étudier l’activité, la reproduction et le mouvement des populations dans un paysage figurent deux grandes catégories : les mesures basées sur la démographie, et les mesures basées sur la variation génétique.

Pour ce qui concerne les approches démographiques permettant d’évaluer la dynamique des populations dans un paysage, les plus utilisées pour les questions spatio-temporelles sont la capture-recapture et la télédétection, qui ne sont pas mutuellement exclusives. La télémétrie par satellite ou par systèmes acoustiques a été très utilisée pour mieux comprendre les mouvements de nombreuses espèces animales, mais elle est généralement restreinte aux espèces de grandes tailles (mais lire Fisher et al. 2021) et à une durée très souvent limitée à une fraction d’une génération (Hooten et al. 2017; Hussey et al. 2015). Un autre exemple de télédétection est constitué par les pièges photographiques qui permettent, par exemple, d’étudier plusieurs espèces simultanément, de suivre des individus particuliers, d’observer des relations trophiques et des comportements, tout ceci de manière peu intrusive (Burton et al. 2015; Caravaggi et al. 2017; Steenweg et al. 2017). Les approches statistiques utilisées pour exécuter ces modèles sont très variées et font souvent appel aux analyses multivariées telles que, par exemple, les analyses canoniques contraintes (Legendre & Legendre 2012) ou les analyses basées sur l’entropie de l’information (Elith et al. 2011).

Pour ce qui concerne les approches génétiques permettant d’évaluer la dynamique des populations dans un paysage, les plus utilisées sont la mesure de la diversité génétique des sous-populations, et la différenciation—ou distance—génétique entre les populations. Ces mesures se basent souvent sur les mêmes données brutes et peuvent donc être utilisées dans une même étude. La diversité génétique peut être utilisée dans les analyses basées sur les nœuds (DiLeo & Wagner 2016; Wagner & Fortin 2013). La nature neutre ou potentiellement adaptative des marqueurs utilisés dicte pour l'étude de quel processus microévolutionnaire (p. ex., flux de gènes vs adaptation locale) les données peuvent être utilisées (Holderegger et al. 2006; Manel & Holderegger 2013). La différenciation génétique peut être utilisée dans les analyses basées sur les liens, mais également pour des résumés locaux de connectivité (DiLeo & Wagner 2016; James et al. 2011a; Wagner & Fortin 2013). La structure génétique, avec assignation d'individus à des groupes génétiquement cohérents, sert souvent de cadre préliminaire aux analyses décrites précédemment (Broquet & Petit 2009). L'ADN environnemental qui est une donnée génétique peut aussi être utilisé pour des analyses de nœuds (Shokralla et al. 2012). Les modèles utilisés pour les analyses de nœuds avec des données génétiques sont similaires à ceux prescrits pour les analyses avec données démographiques. Les modèles utilisés pour les analyses de liens sont plus spécifiques, ce qui a entre autres stimulé le développement d'une sous-discipline qui vise à comprendre comment le paysage facilite ou contraint les processus microévolutionnaires (Holderegger & Wagner 2006; Manel et al. 2003). La télédétection et les estimations génétiques peuvent être complémentaires pour la compréhension de la dynamique d'une espèce. Les différences principales sont que les données génétiques représentent plusieurs générations et qu'elles mesurent le plus souvent des processus post-reproduction dans le cas de la dispersion (Cayuela et al. 2018; Moore et al. 2017).

1.5 Objectifs de la thèse



Comment le paysage influence-t-il la dynamique des populations?

	Chapitre 2	Chapitre 3	Chapitre 4
Menaces	Les perturbations anthropiques altèrent les communautés	Certaines espèces éruptives se répandent et dévastent des écosystèmes Les espèces vulnérables sont affectées par la fragmentation	La stochasticité environnementale peut affecter les populations fragiles
Questions	Comment les organismes utilisent-ils le paysage?	Comment les organismes se déplacent-ils à travers le paysage ?	Quelles sous-populations ont subi un changement majeur au cours du temps ?
Obstacles	De très nombreux processus biotiques et abiotiques façonnent l'utilisation du paysage	Les connaissances à propos de la connectivité sont très limitées pour la plupart des espèces, notamment celles atteignant de nouveaux habitats	La stochasticité environnementale, et ses effets, sont difficiles à observer

Figure 1-4 : Le paysage influence la dynamique des populations qu'il supporte de nombreuses façons. Compléter notre compréhension de l'influence du paysage et de ses perturbations sur l'utilisation des habitats, la connectivité, et les changements temporels de diversité est nécessaire pour mieux faire face à la perte de biodiversité et au changement global. Chaque sous-question sur l'influence du paysage sur la dynamique des populations est pertinente pour la biologie de la conservation, mais présente des défis qui lui sont souvent particuliers.

Développer notre compréhension de la dynamique des populations est nécessaire pour pouvoir limiter la perte de biodiversité et faire face au changement global. Poursuivre l'objectif de mieux comprendre comment le paysage influence la dynamique des populations (**Figure 1-4**) implique de mieux décrire l'effet du paysage sur différents processus affectant la distribution et la connectivité fonctionnelle des espèces. Atteindre cet objectif repose également sur notre capacité à efficacement identifier les changements de biodiversité parfois cryptiques. Remplir cet objectif est d'autant plus urgent et complexe dans les paysages qui n'ont jamais été altérés si profondément et si rapidement. Cependant, des obstacles méthodologiques persistent et freinent les chercheurs et gestionnaires de ressources naturelles dans le progrès vers la réalisation de cet objectif, et ceci dans toutes les facettes de l'étude de l'influence du paysage sur la dynamique des populations. L'application de méthodes innovantes permettant de dépasser ces obstacles est donc étroitement intégrée avec le but de décrypter des patrons spatiaux de biodiversité dans les paysages complexes. C'est pourquoi les trois chapitres d'analyse de cette thèse se penchent sur plusieurs questions centrales à cet objectif, qui sont issues de clairs manques de connaissances écologiques et d'outils pour analyser des données complexes issues de mesures indirectes du paysage, et qui sont stimulées par des besoins urgents en matière de gestion du patrimoine naturel.

Le premier chapitre d'analyse se concentre sur la question générale de comment les organismes utilisent l'espace dans un paysage perturbé. Les perturbations anthropiques directes altèrent la composition et la structure des communautés végétales à tel point que leurs impacts dépassent largement ceux d'autres phénomènes comme le changement climatique (Maxwell et al. 2016). Les effets de la perturbation du paysage tels que la destruction, l'altération et la fragmentation, sur la biodiversité sont généralement bien répertoriés (Fahrig 2003b; Tscharntke et al. 2012). Cependant, les conséquences précises des perturbations sur de nombreuses espèces en interaction, et subissant les mêmes processus spatiaux induits par les perturbations, sont souvent floues. Certaines espèces peuvent bénéficier des perturbations anthropiques directement ou indirectement, alors que d'autres y sont très vulnérables (Devictor et al. 2008; Fisher & Burton 2018; McInturff et al. 2020). Alors que certaines études se concentrent sur les espèces les plus vulnérables et les plus susceptibles d'être les perdants à l'issue d'une perturbation, il est souvent important de considérer l'ensemble de la communauté dans la modélisation de l'effet du paysage sur la dynamique de la communauté pour mieux comprendre, et donc mieux protéger ces espèces (Baselga & Araújo 2009; Hui et al. 2013; Maguire et al. 2016). Dans ce premier chapitre d'analyse, nous tentons d'expliquer comment le paysage influence l'utilisation simultanée de l'espace par de nombreuses espèces de mammifères forestiers, tout en relevant le défi de considérer un grand nombre d'éléments à la fois naturels et anthropiques, à plusieurs échelles, et de garder la communauté intacte dans l'analyse. Les résultats de ce chapitre décrivent le rôle déterminant des perturbations dans le fonctionnement de ce paysage de forêt boréale, et contribuent ainsi à mieux cerner les périls pesant dessus. Il s'agit également d'un exercice visant à mettre en évidence les avancées dans l'incorporation de l'autocorrélation spatiale dans les analyses multivariées, ainsi que leur interprétation (Legendre & Legendre 2012; Peres-Neto & Legendre 2010).

Le deuxième chapitre d'analyse se concentre sur la question générale de comment le paysage affecte le mouvement d'une espèce qui perturbe son biome à grande échelle. La connectivité représente le degré auquel le paysage facilite ou freine le mouvement des organismes (Rudnick et al. 2012; Taylor et al. 1993; Tischendorf & Fahrig 2000). Préserver la connectivité des espèces vulnérables, entraver celle des espèces exotiques envahissantes ou prédire celle des espèces économiquement importantes sont des efforts qui sont de plus en plus considérés comme cruciaux pour la conservation de la biodiversité et la gestion des ressources naturelles (Correa Ayram et al. 2016; Rudnick et al. 2012). Les récentes années ont connu un essor dans l'utilisation de modèles de connectivité fonctionnelle notamment ceux basés sur les données génétiques (Dickson et al. 2019; DiLeo & Wagner 2016; Manel & Holderegger 2013). Malgré tout, il y a beaucoup de défis et limites qui rendent l'utilisation et l'interprétation de ces modèles complexes (Lundgren & Ralph 2019; Peterman & Pope 2021; Richardson et al. 2016; Spear et al. 2010). Dans ce chapitre, nous tentons d'expliquer comment le paysage influence le mouvement d'une espèce d'insecte qui ravage une grande partie de l'écosystème étudié dans le chapitre précédent : la forêt boréale de l'ouest du Canada. En plus d'identifier les éléments du paysage qui contraignent la propagation de l'épidémie actuelle de dendroctone du pin ponderosa dans un nouvel habitat historiquement jamais atteint, nous avons utilisé une méthode puissante qui ne dépend pas de l'opinion des experts pour trouver la relation entre variable environnementale et le coût qu'elle inflige au mouvement des individus (Peterman 2018; Winiarski et al. 2020).

Le troisième chapitre d'analyse se concentre sur la question générale de l'identification de quelle partie du paysage, et donc quelle(s) sous-population(s) s'écarte(nt) de la dynamique des populations typique de ce paysage. Les événements démographiques brutaux peuvent très fortement influencer la variation génétique d'une sous-population. En suivant cette variation génétique dans le temps, il peut être possible de détecter de tels événements modifiant localement la dynamique des populations (Bradburd & Ralph 2019; Fenderson et al. 2020). En intégrant mieux la dimension temporelle aux analyses génétiques, nous pouvons comprendre comment le paysage a historiquement affecté la démographie (Draheim et al. 2018). Cela permet donc de mieux choisir une stratégie de surveillance et de conservation pour les sous-populations qui ont subi un changement génétique atypique, et potentiellement de découvrir l'origine biotique ou abiotique de ce changement a posteriori. Dans ce chapitre, nous construisons et testons une approche permettant de détecter les changements temporels significatifs de diversité génétique. Cette approche est inspirée de méthodes d'écologie des communautés, et est conçue pour des espèces non-modèles et/ou des jeux de données historiques limités. Ce chapitre décrit donc une méthode qui permet de détecter des changements dans la dynamique des populations tels que ceux apportés par les phénomènes étudiés dans les chapitres précédents. Ce chapitre montre également la puissance d'une approche par simulations pour à la fois tester de nouvelles méthodes, mais aussi pour mieux choisir les seuils de significativité afin que chaque étude ait les la possibilité de discerner les compromis entre différents aspects de la performance du test.

La contribution de cette thèse est certes multiple, mais les questions soulevées, les résultats obtenus, et les méthodes mises en valeur convergent en un effort commun. En effet, tous les chapitres visent à mieux comprendre et expliquer les patrons spatiaux liés à la dynamique des populations dans des paysages rendus encore plus complexes par l’action directe ou indirecte de l’homme.

2. CHAPITRE 2 – Modélisation au niveau communautaire de la répartition des mammifères de la forêt boréale dans un paysage de sables bitumineux

Community-level modelling of boreal forest mammal distribution in an oil sands landscape

Julian Wittische[1*], Scott Heckbert [2,3], Patrick M. A. James[1,4], A. Cole Burton[5,6,+], Jason T. Fisher[7,+]

[1] Département de Sciences Biologiques, Université de Montréal, Montréal, QC H3C 3J7, Canada

[2] Department of Geography, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada

[3] Alberta Energy Regulator, Calgary, AB T2P 0R4, Canada

[4] Graduate Department of Forestry, John H. Daniels Faculty of Architecture, Landscape, and Design, University of Toronto, 33 Willcocks St., Toronto, ON, Canada, M5S 2J5

[5] Department of Forest Resources Management, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

[6] Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

[7] School of Environmental Studies, University of Victoria, Victoria, BC V8W 2Y2, Canada

[+] A. Cole Burton and Jason T. Fisher should be considered joint senior author.

Published

Julian Wittische, Scott Heckbert, Patrick M.A. James, A. Cole Burton, Jason T. Fisher. 2021. Community-level modelling of boreal forest mammal distribution in an oil sands landscape, Science of The Total Environment. 755(2): 142500.
https://doi.org/10.1016/j.scitotenv.2020.142500.

Abstract

Anthropogenic landscape disturbances are known to alter, destroy, and fragment habitat, which typically leads to biodiversity loss. The effects of landscape disturbance generally vary among species and depend on the nature of the disturbances, which may interact and result in synergistic effects. Western Canada's oil sands region experiences disturbances from forestry and energy sector activities as well as municipal and transportation infrastructure. The effects of those disturbances on single species have been studied and have been implicated in declines of the boreal woodland caribou (*Rangifer tarandus caribou*). Yet, the specific responses of the mammal community, and of functional groups such as prey and predators, to those interacting disturbances are still poorly known. We investigated the responses of black bear, grey wolf, coyote, fisher, lynx, red fox, American red squirrel, white-tailed deer, moose, caribou, and snowshoe hare to both natural habitat and disturbance associated with anthropogenic features within Alberta's northeast boreal forest. We used a novel community-level modelling framework on three years of camera-trap data collected in an oil sands landscape. This framework allowed us to identify the natural and anthropogenic features which explained the most variation in occurrence frequency among functional groups, as well as compare responses to linear and non-linear anthropogenic disturbance. Occurrence frequency by predators was better explained by anthropogenic features than by natural habitat. Both linear and non-linear anthropogenic features helped explain occurrence frequency by prey and predators, although the effects differed in magnitude and spatial scale. To better conserve boreal biodiversity, management actions should extend beyond a focus on caribou and wolves and aim to restore habitat across a diversity of anthropogenic disturbances and monitor the dynamics of the entire mammal community.

Keywords: Multispecies analysis; Camera trap; Bitumen extraction; Forestry; Cumulative effects; Human footprint

2.1 Introduction

Biodiversity has greatly declined in the Anthropocene (Dirzo et al. 2014), continues to decline despite increasing human response (Butchart et al. 2010), and is unlikely to stop declining in the future given expected trends in human consumption (Bellard et al. 2012). Anthropogenic landscape disturbance is a global driver of this biodiversity loss (Foley et al. 2005) due to its direct effects on habitat availability, quality, and connectivity (Fahrig 2003a). Although many studies have shown that anthropogenic disturbances are major drivers of biodiversity loss, the cumulative and interactive effects of the many forms of disturbances are still poorly known. A key limitation of many previous studies has been a narrow focus on single species, while broader understanding requires approaches that consider community-level responses to disturbance (Cavada et al. 2019; Zipkin et al. 2010).

Forest disturbance affects wildlife species in myriad ways, including spatial legacies relating to the composition and configuration of remaining forest (James et al. 2007). The effect of disturbance on wildlife depends on the character, extent, and intensity of the disturbance, as well as interactions among new disturbances and the legacies of previous disturbances (Côté et al. 2016). Individual and cumulative forest disturbances, and their consequent effects on habitat quality and availability, result in complex wildlife behavioural responses and altered community-level interactions (Courbin et al. 2014). Improved understanding of such mechanisms underlying biodiversity loss is needed to improve conservation decisions (Maxwell et al. 2016). Such decisions are made more difficult by increasing uncertainty when multiple species are threatened (Nicholson & Possingham 2007), such as in hotspots of biodiversity and areas of intensive development.

Western Canada’s oil sands region experiences disturbances from forestry and energy sector activities as well as municipal and transportation infrastructure, which interact to modify the landscape in ways without historical or global analogues (Pickell et al. 2015). These disturbances are also having significant effects on regional vegetation (Abib et al. 2019) and wildlife

(Murdoch et al. 2020), especially mammal communities (Fisher & Burton 2018). Notably, boreal woodland caribou (*Rangifer tarandus caribou*) populations, an ecotype of caribou, are declining across Canada partly due to industrial disturbance, and the species has consequently been listed as "Threatened" in the Species at Risk Act (Government of Canada). Caribou declines across its range have been driven by industrial development and consequent habitat loss, predation, and indirect processes such as apparent competition (Beauchesne et al. 2014; Boutin et al. 2012; Hebblewhite 2017; Wasser et al. 2011).

Disturbance-mediated changes to the spatial structure of boreal forest landscapes have altered historical relationships between caribou, their predators, and other prey species. Predation of adult and neonate caribou by wolves (*Canis lupus*) (Latham et al. 2011b; McLoughlin et al. 2003) and of neonates by black bears (*Ursus americanus*) (Latham et al. 2011a) is thought to be the primary proximate cause of decline for caribou in the oil sands region of northeast Alberta, as well as in other parts of its range (Festa-Bianchet et al. 2011). Wolf numbers have increased, following increases in populations of white-tailed deer (*Odocoileus virginianus*), and possibly moose (*Alces alces americanus*) and beavers (*Castor canadensis*) (Latham et al. 2011c, 2013a). This increase in prey has likely followed the landscape alteration associated with human development, through the increased ungulate forage provided by the conversion of mature forest into early seral vegetation (Fisher & Wilkinson 2005). Anthropogenic linear features often permeate exploited forest ecosystems. They generally include a range of features from trails to roads and forestry electrical transmission lines, with the addition of exploration seismic cutlines (simple and 3D) and pipelines in oil sands areas where petroleum is extracted. The intensity of linear features use by wolf (Dickie et al. 2016) and white-tailed deer (Fisher et al. 2020) are strongly linked, and affected by both human use and obstacles created for attempts at restoration of linear features (Keim et al. 2019; Tattersall et al. 2020a). Collectively, previous research suggests that cumulative effects of multiple, linear and non-linear, disturbances are effecting changes in community structure and predator-prey interactions (which most prominently, drive caribou population declines). However a community approach to mammal responses to widespread landscape change has only begun to be characterized (e.g., Burgar et al. 2018).

Further community-level research is needed to advance our understanding of how community dynamics are changing within changing landscapes.

In this research, we address the knowledge gaps on community-level response to disturbance by: 1) comparing the effects of natural habitat vs. anthropogenic disturbance on predators, prey and the whole mammal community; 2) comparing the effects of linear vs. non-linear anthropogenic disturbances; 3) evaluating the influence of species interactions relative to habitat and disturbance. We suspect that simultaneous consideration of the responses of multiple interacting species may lead to different conservation recommendations than the current focus on only one or a few focal species, such as caribou and wolves.

Although several single-species studies have examined behavioural response to one type of anthropogenic feature—such as wolves moving faster down seismic lines (Dickie et al. 2016)—to date there has been no examination of responses to multiple anthropogenic disturbance types across the community. Are the effects of disturbance merely reflective of a single species' niche and confined to behaviour, or are they consistent across multiple species, reflective of a species' functional role (e.g., predator vs. prey), and do they scale up to occurrence frequency? This is the novel perspective of a community analysis approach seeking to find generalizable patterns across suites of mammal species. We hypothesized that anthropogenic features would influence predators to a greater extent than prey, given the documented heavy use of anthropogenic landscape features by wolves (Latham et al. 2011a; Lesmerises et al. 2012; McKenzie et al. 2012; Wasser et al. 2011) and the principle of trophic amplification (Kirby & Beaugrand 2009). We also hypothesized that among anthropogenic features, linear features would predominate as the best-supported predictors of predators' occurrence frequency, given their apparently outsized influence in the single-species literature. Finally, we hypothesized that prey with the highest biomass in this system – snowshoe hares and white-tailed deer – would most influence predators' occurrence frequency.

To test these hypotheses, we measured the frequency of occurrence of caribou, moose, wolves, bears, and other species, relative to both natural habitat and disturbance associated with anthropogenic features within Alberta's northeast boreal forest. We use a novel community-based methodological framework based on count data from unmarked populations which enables us to explore how varying mammal habitat use, as measured by frequency of occurrence across space, is explained by different types of predictors in the landscape. We used variation partitioning (Borcard et al. 1992) to make the most of those community-based constrained ordinations. More precisely, we partitioned the explanatory power (R^2) of community-level models of the occurrence frequency of 11 mammal species between best-supported natural and anthropogenic predictors. Building on the fractions of variation explained by different groups of predictors, we discuss how our study provides ecological insights beyond studies focused on a single species. Based on our findings, we stress the conservation value of the complementary information brought by a community-level modelling approach to the boreal mammal community.

2.2 Materials and methods

2.2.1 Study area

Our study was conducted in the boreal forest northeast of Lac La Biche, Alberta, Canada (**Figure 2-1**). The study area included the area around Christina Lake and Winefred Lake, north of the Cold Lake Air Weapons Range, and is approximately 3 000 km^2. This mosaic-like forested landscape is mostly composed of white (*Picea glauca*) and black spruce (*Picea mariana*), aspen (*Populus tremulodies*), jack pine (*Pinus banksiana*), and muskeg dominated by bog Labrador tea (*Rhododendron groenlandicum*). Anthropogenic disturbance is extensive with forestry and energy sector activities, including roads distributed over the study area (**Figure 2-1**). Averaged over 1-km^2 cells within the study area, the percentage of seismic lines area per cell is 1.02 [SD 0.89] %, and the density of total disturbance area per cell is 2.72 [SD 3.27] %.

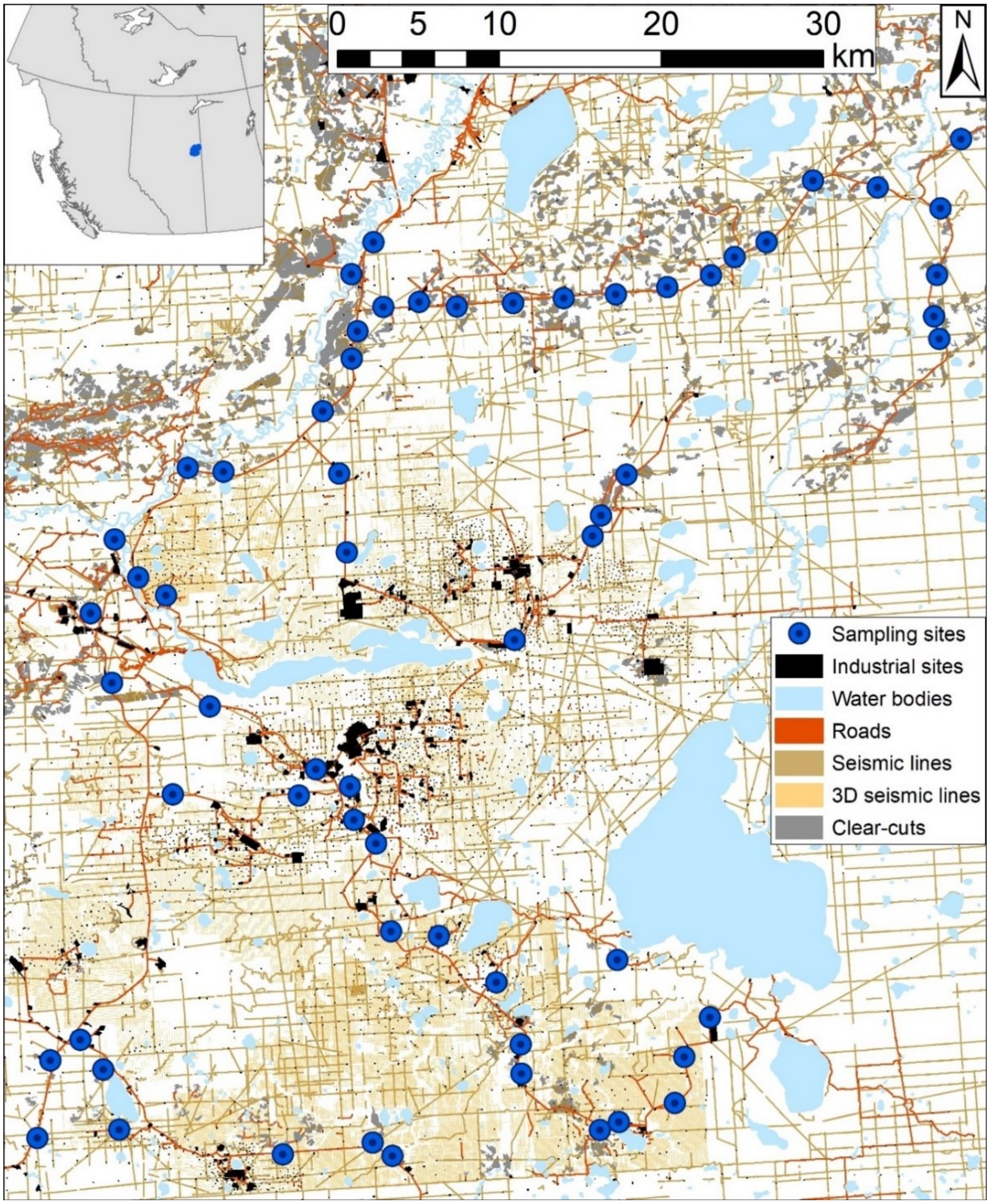


Figure 2-1 : The 61 camera-trap sites (dark blue circles) where mammal occurrence was surveyed in the northeast boreal forest near Winefred Lake, Alberta, Canada, between October 2011 and October 2014. Please note that the grid pattern associated with 3D seismic lines is at such a fine scale that individual lines resemble a fine mesh on our maps.

2.2.2 *Faunal data*

Mammal frequency of occurrence was sampled at 61 camera trap sites (**Figure 2-1**). Reconyx PC900 Hyperfire^{TM} infra-red remote digital cameras (Holmen, WI, USA) were deployed for three years: October 2011–October 2014. The distribution of habitat types that are sampled within a landscape can greatly influence landscape-scale inferences (Dale & Fortin 2014; Leroux et al. 2007). To increase the reliability of our extrapolations to other landscapes, and therefore have more general results, the camera traps were distributed across the landscape by using a constrained random stratified design intended to equalize representation of all habitat types. Camera locations were selected from among a set of 1-km resolution raster cells (our statistical units) covering the study area. We imposed a minimum 2-km distance between cameras to meet model assumptions of independence among statistical units (see Fisher & Burton 2018 for more details). Among all 1-km^2 cells from the 3 000 km^2 study area, 406 cells were identified as being reliably accessible; 61 final candidate cells were selected randomly within several forest strata (see *Environmental Data*). Cells that had no access were dropped and replaced from this pool. This had the effect that roads and trails were more highly represented within 250-m radius of sampling sites than proportional in the landscape. However, because these linear features were so ubiquitous in this landscape (**Figure 2-1**), this effect disappeared beyond this distance, wherein linear features were represented in our sample proportional to the landscape.

Within grid cells, cameras were systematically deployed on active wildlife trails, that is trails with evidence of recent and substantial wildlife travel, at a minimum distance of 200-m from roads and human trails, including the ones used to access the camera trapping sites. Each trail was independent from others in adjacent cells. The mean and median distances between sampling sites were 33.1 [SD 17.3] km and 32.2 km, respectively. The resulting design captured the range of this landscape's anthropogenic footprint and natural heterogeneity. We obtained permission from government land officers and industry leaseholders to access all sampling areas. Camera traps were not lured nor baited. The camera trap survey had a total sampling effort of 60,937 camera trap-days which yielded 164,519 photographs, providing 141,140 images of mammal species. Among those 141,140 images, 134,482 were identified to species: black bears (2,657 images), grey wolves (2,508 images), coyotes (*Canis latrans*; 2,290 images), fishers (*Pekania pennanti*; 326 images), lynx (*Lynx canadensis*; 1,940 images), red fox (*Vulpes vulpes*; 197 images), American red squirrel (*Tamiasciurus hudsonicus*; 491 images), white-tailed deer (112,648 images), moose (500 images), caribou (273 images), and snowshoe hare (*Lepus americanus*; 10,652 images). The final dataset therefore included six predator species and five prey species.

To track responses of these mammals to landscape features, we used an index of occurrence frequency based on a minimum interval of one hour between consecutive detection events of the same species at the same site. Such corrected occurrence frequency indices are commonly used in camera trap studies (Burton et al. 2015; Keim et al. 2019; Steenweg et al. 2017; Tattersall et al. 2020a). Occurrence frequency metrics from camera trap surveys incorporate both the number of individuals at a site and the repeated use of a site by an individual(s), which is mediated by movement behaviour (Stewart et al. 2018). We temporally truncated data to avoid the problem of curious, lingering species such as bears being over-represented in the occurrence frequency metric because they take time to investigate the camera, whereas deer and other species do not. The choice of one hour was also guided by an assessment of the maximum time a camera-exploration behaviour occurred in our dataset. We further consider this metric to be reliable as recent simulation-based work has shown a close relationship between abundance and detection rates for multiple species with varying home range sizes and movement behaviour (Broadley et al. 2019). Image captures of groups of animals were recorded as a single occurrence. Herein we study co-occurrences, which we conceptually link to the potential for trophic interactions (e.g., predator and prey). As the field settings were the same for all cameras and because we do not compare occurrence frequency across species, we avoided introducing a systematic bias in our analysis through camera placement.

We standardized for sampling effort by dividing the number of detections by the number of camera-days at each site. Camera sampled for an average of 997.44 [SD 152.88] days. The resulting site × species matrix of occurrence frequency was standardized using a Hellinger transformation prior to analysis, which is a common way to improve the statistical properties of redundancy analysis (RDA) models, while respecting the role of rare species (Legendre & De Cáceres 2013; Legendre & Gallagher 2001; Peres-Neto et al. 2006).

2.2.3 Environmental data

We quantified natural landscape composition using the Alberta Vegetation Inventory (AVI), a digital, vector format, species-level forest inventory dataset provided by the Government of Alberta (**Table 2-1**). These data were reclassified into two broad categories. We classified black spruce (*Picea mariana*), jack pine (*Pinus banksiana*), white spruce (*Picea glauca*), and balsam fir (*Abies balsamea*) as "coniferous". Paper birch (*Betula papyrifera*), aspen (*Populus tremuloides*), balsam poplar (*Populus balsamifera*) and tamarack (*Larix laricina*) were classified as "deciduous". Within each 1-km^2 raster cell, we extracted the area of each AVI polygon along with the dominant canopy species in the polygon and multiplied that by the percentage of the canopy in that polygon. We classified each polygon as lowland if the moisture regime was recorded as aquatic or wet, or upland if it was not. The canopy cover designations and moisture regime information were combined to create thirteen land cover categories (**Table 2-1**). We calculated the area (m^2) of each habitat category around each sampling site at different scales: we used 20 radii from 250 m to 5000 m, in 250 m increments. Those scales have previously been shown to be meaningful for the boreal forest mammal community (Fisher et al. 2011).

Table 2-1 : Description and classification of natural and anthropogenic data extracted in our study area and used in the selection process.

Habitat or Feature class	Source	Description
Upland deciduous	AVI	(Aw, Pb, Bw >=70% canopy), moisture = d or m
Lowland deciduous	AVI	(Aw, Pb, Bw >=70% canopy), moisture = w or a
Upland mixedwood	AVI	(40% -60%) canopy, moisture = d or m
Lowland mixedwood	AVI	(40% -60%) canopy, moisture = w or a
Upland spruce	AVI	(Sb, Sw, Fb >=70% canopy), moisture = d or m
Lowland spruce	AVI	(Sb,Sw,Fb >=70% canopy), moisture = w or a
Pine	AVI	All Pj (>=70%)
Tamarack	AVI	All Lt (>=70%)
Open wetland	AVI	<6% crown closure; moisture = w or a
Upland shrubs	AVI	>25% shrub cover; <6% tree cover; moisture = d or m
Water	AVI	Standing or flowing water
Nonforest	AVI	Areas with < 6% canopy
Cutblock	ABMI	Forest harvested cutblocks of any age
Block feature	ABMI	Combination variable including mining borrow pits, dugouts, sumps, industrial sites, and "other disturbed vegetation". No mature trees.
Well site	ABMI	Energy sector sites including a well and surrounding area, usually grassy vegetation.
3D seismic line	UALF	3D seismic energy sector exploration lines deployed in a high-density hashmark pattern.
Cutline	UALF	Traditional single energy sector exploration lines, less dense than 3D seismic lines.
Roads (all)	UALF	Combination variable of one- and two-lane roads, gravel and paved, and unimproved roads.

Table 2-2 (continued): Description and classification of natural and anthropogenic data extracted in our study area and used in the selection process.

Habitat or Feature class	Source	Description
Electrical transmission line	UALF	
Trails (all)	UALF	Combination of trails and truck trails.
Driveway	UALF	
One-lane gravel road	UALF	
One-lane paved road	UALF	
Pipeline	UALF	Energy sector pipelines and their rights of way, usually wide and grass-covered.
Rail	UALF	
Trail	UALF	Trails navigable by off-road vehicles, horses, and people but not trucks.
Truck trail	UALF	Trails navigable by trucks but without gravel.
Two-lane gravel road	UALF	
Two-lane paved road	UALF	
Unimproved road	UALF	
Borrow/Pit/Dugout/Sump	ABMI	
Industrial site	ABMI	
Other disturbed vegetation	ABMI	
Rail (hard surface)	UALF	
Rail (vegetated verge)	UALF	
Road (hard surface)	UALF	
Road (vegetated verge)	UALF	
Road/Trail (vegetated verge)	UALF	
Seismic line	UALF	

[1] AVI – Alberta Vegetation Index; UALF = University of Alberta Linear Features Map Updated 2012; ABMI = Alberta Biodiversity Monitoring Institute Human Footprint Map Updated 2010.

[2] Aw = aspen, Pb = poplar, Bw = white birch, Sb = black spruce, Sw = white spruce, Fb = balsam fir, Pj = jack pine, Lt = tamarack/larch; d = dry, m = mesic, w = wet, a = aquatic

[3] Linear features were buffered for areal calculations as follows: 'Two-Lane Undivided Paved Road': 9m; 'One-Lane Undivided Paved Road':6m; 'Rail Line': 5.5m; 'Rail Line- spur': 5.5m; 'Two-Lane Gravel Road': 7m; 'One-Lane Gravel Road': 5m; 'Driveway': 2m; 'Unimproved Road': 6m; 'Trail': 6m; 'Truck Trail': 6m; 'Electrical Transmission Line': 17m; 'Pipeline': 12m; '3D': 2m; 'Cutline': 2m.

We calculated the area (m^2) of each anthropogenic feature around each sampling site (**Table *2-1***) using the Alberta Biodiversity Monitoring Institute Human Footprint Map Version 1.1 (ABMI 2010) in the same way we did for natural habitat categories, that is using 20 radii from 250 m to 5000 m, in 250 m increments. Areas of each environmental variable within the surfaces delimited by the radii we picked, can be understood as proportions of the neighbouring landscape belonging to that variable. As low-impact seismic lines were not sufficiently captured in the ABMI (2010) dataset, we used buffered linear features data from 2012 (ABMI, unpublished) to supplement the ABMI (2010) data (**Table 2-1**). We calculated the percentage of the area occupied by linear features (buffered to create polygons from polylines) around each sampling site. Similar linear features open data is now also available through open governmental websites (Government of Alberta 2019).

2.2.4 Creation of multi-scale spatial variables

Spatial autocorrelation in ecological data can confound efforts to characterize the effect of the environment on community structure (Dale & Fortin 2014). Such autocorrelation can be included and controlled for in statistical models of ecological variation through the use of synthetic spatial variables such as Moran’s eigenvector maps (MEM) (Borcard & Legendre 2002; Borcard et al. 2004; Dray et al. 2006; Legendre & Legendre 2012). MEMs are computed as the eigenvectors of a spatial weighting matrix based on a set of geographical coordinates (i.e., sample sites). We used the *dbmem()* function from the *adespatial* 0.3-7 R (R Core Team 2019) package (Dray et al. 2019) to create the MEMs. MEM analysis is a multi-scale spatial analysis that produces orthogonal spatial eigenfunctions, which we can use to identify spatial patterns of variation in the response variable(s) across a range of spatial scales. The first few MEM variables created through the algorithm (large positive eigenvalues, which correspond to large Moran's I coefficients) represent broad scale processes. MEM variables with smaller eigenvalues, which correspond to smaller Moran's I coefficients, model fine spatial autocorrelation generated by local processes. (**Supp. Figure 2-1**). MEMs can represent either positive or negative spatial autocorrelation.

2.2.5 Regression framework

We used RDA (Legendre & Legendre 2012) to model the association between occurrence frequency of mammal species and groups of environmental and spatial predictors. We used the *rda()* function from the *vegan* 2.5-2 R package (Oksanen et al. 2018) to perform the RDAs. Groups of predictors (submodels) were used to model hypothesized forces driving mammal abundance, including spatial autocorrelation (MEMs), anthropogenic disturbance, natural habitat, and predation, when relevant. The response term of the models was multivariate and included the occurrence frequency of several species assembled in a data matrix, and therefore did not consist of one variable with occurrence frequency of different species lumped together. We chose a community-level approach because it explicitly addresses processes creating co-occurrence patterns beyond shared abiotic requirements of species, such as biotic interactions (Baselga & Araújo 2009; Maguire et al. 2016). Additionally, community-level models have been shown to be at least similar to an assemblage of single-species models in predicting species distributions, with significant improvement for rare species (Hui et al. 2013; Maguire et al. 2016).

We used a two-step approach to narrow down the relevant predictors for each analysis: 1) selection of the best scales for each environmental predictor and 2) selection of the best environmental predictors overall. First, we used a modified forward selection procedure (Blanchet et al. 2008) to select the best scale(s) for each environmental predictor. We used the *forward.sel()* function from the *adespatial* 0.3-7 R package (Dray et al. 2019) with a alpha threshold of 0.1, instead of the threshold of 0.05 used in the rest of our analyses, to be more liberal during this first step. A potential pitfall associated with the use of forward selection to select environmental predictors is the overestimation of the amount of explained variance (Diehr & Hoflin 1974). The modified two-step forward selection we used avoids overestimating the coefficient of multiple determination by conducting a global test of significance prior to using two stop criteria (Blanchet et al. 2008). We tested the significance of the joint effect of the preselected variables on response variables within each sub-model independently using a permutation test (Legendre & Legendre 2012). We tested the global significance separately for negative and positive spatial eigenvectors to provide enough degrees of freedom (Blanchet et al. 2008) and therefore applied Sidak's correction to the alpha level of rejection of the null hypothesis (Sidak 1967) in order to have an appropriate rejection rate overall.

Second, if a sub-model was deemed significant in the previous step, we used the same forward selection procedure used in the first step to select the best variables to keep in a final model. When using MEM among the spatial predictors, we tested and forward selected separately for each species and used the union of all selected predictors (Peres-Neto & Legendre 2010). To be clear, we did not select variables and radius sizes per species but simultaneously for all species, neither did we select the same radius for all environmental variables. By varying the radii used to calculate environmental characteristics, we took into consideration varying species requirements for habitat use.

Finally, we applied variation partitioning to our selected multivariate model of mammal occurrence frequency. Using this approach, we sought to quantify the unique and shared contributions of each retained predictor to the variation of community composition data (Borcard et al. 1992; Griffith & Peres-Neto 2006). Specifically, we assessed the unique and shared contributions of the following categories of predictors: 1) Natural habitat, 2) Anthropogenic features, 3) Spatial autocorrelation, and 4) Occurrence frequency of predators or prey when relevant. Variance partitioning was implemented using the *varpart()* function in the *vegan* 2.5-2 R package (Oksanen et al. 2018) to perform the variation partitioning. Therefore, the intersecting fractions of the variation explained by two (or more) submodels describe the variation that is not distinguishably explained by one of them and do not describe the effect of interactions between submodels. We also used variation partitioning on the following submodels: 1) Linear anthropogenic features, and 2) Non-linear anthropogenic features in a subsequent analysis.

2.3 **Results**

2.3.1 Performance of models and submodels

The frequency of occurrence of the mammal community across this highly disturbed landscape was shaped by a combination of anthropogenic features, natural habitat, spatial autocorrelation, and predation, as all four of these submodels were consistently significant in explaining variation in the observed occurrence frequency of mammals (**Table 2-3**).

As hypothesized, the anthropogenic features submodel explained most (30% with a total of 37% with a full model) of the variation in the occurrence frequency of predators. On the other hand, the natural habitat submodel explained most (43% with a total of 51% with a full model) of the variation in the occurrence frequency of prey (**Table 2-3**). The natural habitat submodel also explained most of the variation at the whole community level, that is when considering all species in the dataset.

Table 2-3 : Adjusted coefficients of determination for each submodel, along with the selected variables, for each functional group. Numbers in parentheses indicate scale of selected variables in kilometres. Variables selected across functional groups are in bold.

FUNCTIONAL GROUP		
Predators	Prey	Whole community
SUBMODEL		
Natural habitat		
0.17 **Upland deciduous (1.75)** Lowland mixed (5) Lowland mixed (1.75)	0.43 **Upland deciduous (0.75)** Lowland spruce (0.25) Open wetland (0.25)	0.33 **Upland deciduous (0.75)** Lowland spruce (0.25) Open wetland (0.25)
Anthropogenic features		
0.30 **Block feature (0.5)** **3D seismic line (0.5)** **Road/Trail (veg.) (1.5)** Industrial site (4) Industrial site (5) Road (veg.) (4.75) Trails (all) (2.5)	0.27 **Block feature (1)** **3D seismic line (0.25)** **Road/Trail (veg.) (1.5)** Cutblock (0.25)	0.22 **Block feature (1)** **3D seismic line (0.25)** **Road/Trail (veg.) (1.5)** Cutblock (0.25)
Spatial variables		
0.14 Longitude MEM 2 MEM 3 MEM 6	0.07 MEM 8	0.04 MEM 8
Other functional group		
0.10 Snowshoe hare	0.20 Coyote Lynx	not applicable
All submodels		
0.37	0.51	0.39

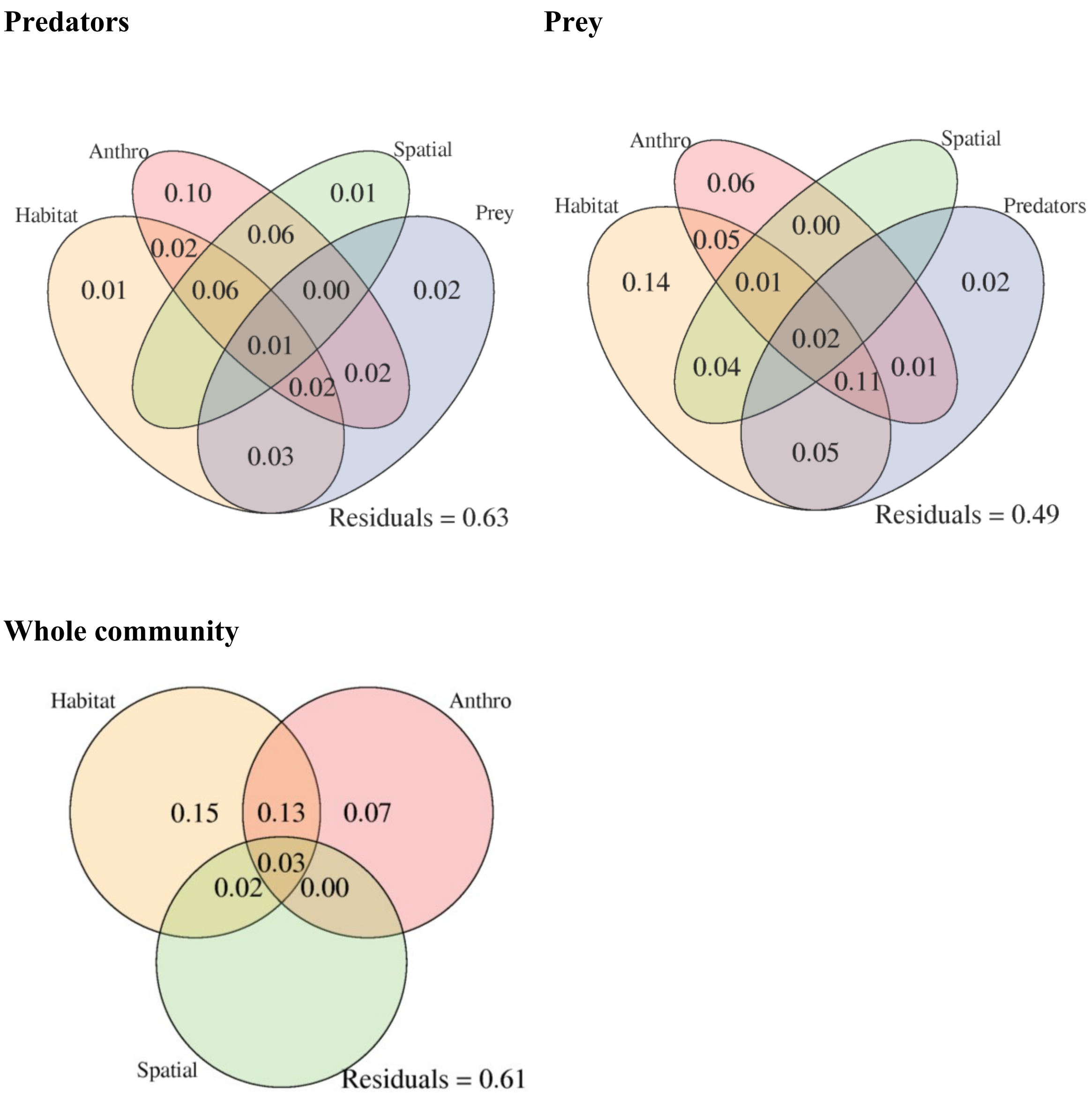


Figure 2-2 : Partition of the variation of the occurrence frequency of predators, prey and the whole community, among the explanatory submodels. Fractions of the explained variation are expressed in adjusted coefficient of determination. Intersections represent parts of the variation which prediction is shared by different submodels. "Habitat" refers to natural habitat and "Anthro" refers to anthropogenic disturbances. Values less than or equal to 0 are not shown.

Although the anthropogenic features submodel was the best submodel for predators, it explained about the same amount of variation in both predators (30%) and prey (27%) (**Table 2-3**). Although significant, the spatial submodel explained the least variation in all three (predators, prey and whole community) analyses (**Table 2-3**) and most of the variation explained by the spatial submodel was consistently shared with at least one other submodel (**Figure 2-2**). This is expected as natural and anthropogenic predictors were themselves spatially-structured, at a scale close to one of the scales covered by MEM. Using a functional group (prey or predators) as a submodel consistently explained more than 10% of the variation in occurrence frequency (**Table 2-3**).

Pure fractions of explained variation — those that were unique to submodels — were consistently smaller than the sum of shared/common fractions within a submodel (**Figure 2-2**). In ecological terms, this means that for example, natural habitat and anthropogenic features — which are two different submodels — jointly explain mammal occurrence frequency. That is their respective explanatory power to describe mammal occurrence frequency intersect substantially (Legendre & Legendre 2012). High shared fractions also mean that only a limited part of the variation in occurrence frequency is explained by only one of the submodels. However, the best submodel for prey (natural habitat), predators (anthropogenic features), and the whole community (natural habitat) consistently had substantial pure fractions (**Figure 2-2**). General qualitative conclusions were not meaningfully affected by changing the temporal bin size (1-hour delay) of our index.

2.3.2 Most important environmental variables and canonical axes

Anthropogenic features selected as important — whether for predators, prey or the whole community — included a diversity of linear features (3D seismic lines, roads, and trails), as well as polygonal disturbance features (**Table 2-3**). Upland deciduous habitat was the only habitat identified as important across functional groups, although lowland mixed forest habitat, open wetland and lowland spruce were also selected by prey or predators (**Table 2-3**). Snowshoe hare was the only prey species selected as a significant predictor of predator occurrence frequency (**Table 2-3**, **Table 2-3**). This finding only partially supports our hypothesis. The occurrence frequency of lynx and coyote best explained the occurrence frequency of prey, which we had not hypothesized based on the literature (**Table 2-3**). The vast majority of the explained variation is concentrated on the first axis for prey (**Supp. Table 2-2**) and the whole community (**Supp. Table 2-3**) with respectively 95% and 87% of the explained variation explained by the first RDA axis, compared with 50% for predators. Coyotes and lynx have the highest fit among predators in the whole community analysis (**Supp. Table 2-3**) which parallels their selection as predictors for prey occurrence frequency (**Table 2-3**). In **Supp. Table 2-1**, wolf occurrence frequency is not explained by the same canonical axis (RDA 2) as black bear or coyote occurrence frequency (RDA 1), with RDA 1 explaining more variation than RDA 2. Triplot visualizations of the RDA results also highlight the differences between species and their relationships with the environmental variables selected previously (**Supp. Figure 2-2**).

2.3.3 Spatial scale

Natural habitat influenced the occurrence frequency of prey at smaller spatial scales than that of predators, as all selected variables were best supported at scales under 1 km for the occurrence frequency of prey, whereas there was no selected variable with a scale below 1.5 km for the occurrence frequency of predators, with one variable selected at the largest scale available (5 km). Linear and block anthropogenic features which were selected for the occurrence frequency of predators were supported from smaller to larger scales, whereas anthropogenic features were consistently selected for the occurrence frequency of prey at smaller scales (**Table 2-3**).

The same pattern was supported by the selected variables in the spatial submodel. Indeed, broader scale MEMs (e.g., MEM 2 which may represent landscape-wide variation) and intermediate scale MEMs (e.g., MEM 6 which may represent more local between-site variation) were selected for predators, while the only MEM selected for prey (MEM 8) was a smaller scale MEM (**Table 2-3**).

2.3.4 Linear vs. Block features

Within the anthropogenic features submodel, the relative importance of linear and non-linear features varied across prey and predators and as hypothesized, linear features seemed to be more important than block features in explaining predator occurrence frequency (**Table 2-4**). Indeed, linear features explained twice as much variation as polygonal features do for predators, while both types of anthropogenic features explained around the same amount of variation for prey (**Table 2-4**). Moreover, in contrast to the analyses using the natural habitat, spatial and predation submodels, there was very little shared explained variation among linear and block features, with most of the explained variation being in the pure fractions.

Table 2-4 : Variation partitioning of occurrence frequencies among linear and block features, within the Anthropogenic features submodel. Pure and shared fractions of the explained variation are displayed.

	FUNCTIONAL GROUP	
	Predators	Prey
FEATURE TYPE		
Linear (pure)	0.20	0.12
Shared	0.01	0.02
Block (pure)	0.09	0.13

2.4 **Discussion**

2.4.1 Impacts of anthropogenic features on the boreal mammal community

Anthropogenic landscape disturbance has a major influence on the boreal oil sands mammal community. We found that the anthropogenic disturbances such as cutblocks and linear features such as 3D seismic lines, roads, and trails influenced both predator and prey communities (**Table 2-3**, **Table 2-3**). These features had a greater effect on predators than prey to the extent that anthropogenic legacies were found to have a greater effect on predator occurrence frequency than natural habitat variation. The variance explained by anthropogenic features was distinct from that explained by natural habitat for predators. We also found that anthropogenic features affect predators at larger spatial scales than for prey (**Table 2-3**). Our findings greatly elaborate the relationships between anthropogenic disturbance and mammal occurrence frequency (Fisher & Burton 2018; Toews et al. 2017a, 2018), and further contribute to our understanding of how the persisting spatial legacies of industrial forestry and the energy sector affect different components of the boreal mammal community.

The mechanism by which human disturbance affects mammal communities is the conversion of intact mature forest to early seral stages, changing both resource availability and animal movement. Early seral vegetation in young cutblocks is preferred by moose (Cederlund & Okarma 1988; Rempel et al. 1997). Clearcutting can produce spatial legacies similar to those created by wildfire in terms of spatial extent and successional responses which results in an increase in available forage in the years following disturbance (Crête et al. 1995; Krefting 1974). The early seral vegetation present in cutblocks is also beneficial for white-tailed deer (Fisher & Wilkinson 2005; Fisher et al. 2016; St-Louis et al. 2000). Block features (e.g., clearcuts) were identified as important for predators, which is not surprising given their higher abundance of anthrophilic and early-seral-dependent prey such as white-tailed deer (Fisher & Burton 2018; Fisher et al. 2020). Our results are also consistent with the hypothesis that linear features facilitate predators including coyotes (Boisjoly et al. 2010a) and wolves (Dickie et al. 2016;

McKenzie et al. 2012; Whittington et al. 2011). Whether the associations we highlighted are behavioural, numerical (and associated with resource competition, apparent competition, or human-driven increased mortality), or both, requires further study.

Considering the nature of the verges of roads and trails is relevant to understanding their effects on the boreal mammal community. By using many variables related to roads and trails in the selection process (**Table 2-1**) instead of grouping them a priori, we allowed the explicit modelling of roads and trails with vegetated verges (**Table 2-1**). Our results suggest that potentially negative effects of roads and trails associated with direct mortality or risk avoidance may be offset by the presence of vegetated verges on the sides of those linear features which provide forage subsidy. Roadside forage subsidy has been reported to be attractive to white-tailed deer (Bellis & Graves 1971). Furthermore, telemetry studies have suggested such a mechanism is possible for white-tailed deer (Darlington 2018).

Although anthropogenic disturbance explained about the same amount of variation in prey and predator communities, this variation differed in its influence relative to natural habitat. More than 70% of variation in prey occurrence explained by anthropogenic features was shared with natural habitat (**Figure 2-2**); this means variation in mammal occurrence frequency can be explained by either. The fact that anthropogenic features shared such a large part of the variance they explain with that of natural habitat might be a clue that the effect of anthropogenic features on prey is more dependent on the type and/or density of surrounding natural habitats than for predators. Indeed, mammals, and prey in particular, require a diversity of boreal forest habitats to live and may conditionally avoid certain natural and altered habitats. For example, moose require mature forest for cover (Forbes & Theberge 1993; Herfindal et al. 2009; Månsson et al. 2007) and its availability could supersede any positive effects of forage subsidy from vegetated linear features or industrial sites. Smaller prey (e.g., snowshoe hare, red squirrel) are known to avoid recent clearcuts (Fisher & Wilkinson 2005; Lewis et al. 2011; Newbury & Simon 2005) and linear features (Bakker & Van Vuren 2004; Fisher & Wilkinson 2005; Oxley et al. 1974). This appears to be a major difference between prey and predators in our study area, where only about

33% of the explained variation in predator abundances was shared between anthropogenic features and natural habitat. However, further research is needed to identify which parts of the natural landscape interact with anthropogenic features. If interactions are indeed responsible for the patterns of shared variation in our data, then an increase of the density of anthropogenic features could change the dynamics of predators more independently of surrounding natural habitat than for prey.

Both anthropogenic features and spatial structure (i.e., MEMs) were selected at small scales by prey. This suggests that when linear and block features are physically proximal, their effect on prey is strongest. A heavily disturbed landscape would mean that strong but small-scale effects on prey are nonetheless widespread. Because predators are affected by environmental processes at larger scales (i.e., farther away), and because predation plays a major role in prey distribution and hence community composition, the effect of anthropogenic features on predators may also translate on prey and vice versa (Wisz et al. 2013).

2.4.2 Linear vs. Block features

Both linear and block features influence mammal occurrence frequency in the boreal forest, although their effects differed between predator and prey species. As hypothesized linear features were more important than block features in explaining predator occurrence (**Table 2-4**). Nonetheless, block features still explained about 10% of the variation in predator occurrence. Those 10% represent 25% of the total variation explained by the full model. Block and linear features explained about the same amount of variation in prey species. These contrasting results illustrate the importance of how the shape of human disturbance (linear vs. block) affects mammal communities – a basic tenet of landscape ecology but often neglected in management. Previous studies have shown the importance of block features for specific boreal species, such as the avoidance of clearcuts by caribou (Lafontaine et al. 2019; Sorensen et al. 2008). Other studies compared both linear and block features on specific species (e.g., Houle et al. 2010;

Beauchesne et al. 2013), however, to our knowledge, this is the first study that explicitly compared the two types of features from a community perspective.

A surprising outcome of partitioning variation explained by linear vs. block features was that each explained different parts of the variation in occurrence frequency, in both prey and predators. Most of the explained variation is present in pure fractions (**Figure 2-2**); linear features and block features affect the whole mammal community in measurably different ways. The low shared fraction does not mean there is no interaction between these feature types, but that they explained very different parts of the variation in mammal occurrence frequency (Legendre & Legendre 2012). Focusing only on linear features as an ecological mechanism yields only part of the answer; and setting only linear features as a target for conservation will solve only a part of the problem.

2.4.3 A fuller picture of the community dynamics

While a large amount of research has focused on wolves and black bears as the main predators in this community, lynx and coyotes are also important components. Indeed, their occurrence frequency significantly influenced the occurrence frequency of prey (**Table 2-3**). This could be expected given their high densities in boreal oil sands landscapes (Burgar et al. 2019). Coyotes are adaptable to disturbance and have been expanding their range for several centuries (Laliberte & Ripple 2004; Levy 2012) and they can markedly affect mammal communities (Heim et al. 2017). Large prey (e.g., deer) tend to represent the majority of coyote diet in eastern Quebec (Boisjoly et al. 2010a). Large coyotes have successfully preyed on moose (Benson & Patterson 2013) and caribou calves (Crête & Desrosiers 1995; Latham et al. 2013a) with a detrimental effect on recruitment (Frenette et al. 2020; Lewis et al. 2017). Wolves have weaker environmental associations than coyotes (**Supp. Table 2-1**, **Supp. Table 2-3**, and **Supp. Figure 2-2**). This does not mean that wolves are not influenced by the landscape, but perhaps suggests that the frequency of occurrence of wolves could change less with further anthropogenic perturbations of the landscape than that of coyotes. Futureproofing conservation decisions will require consideration of coyote's responses to disturbance, as well as wolves'.

2.4.4 Study limitations and further research

Choosing adequate indices of occurrence frequency/relative abundance, or other demographic metrics, from unmarked data, such as most camera trap data, is challenging (Burton et al. 2015; Chandler & Royle 2013; Dénes et al. 2015). Here we assumed that zeros observations were true zeroes, rather than detection error, which we limited by avoiding obstructions, limiting false triggers, and optimizing wildlife detection, but we recognize that habitat structure and animal movement both influence detection. Thoroughly validating the interpretation of occurrence frequency/relative abundance indices as proxies of true abundance, or other demographic properties, involves testing assumptions about the functional relationship between the demographic property under study and those indices. Indeed, many assumptions behind the metrics based on camera trap data remain untested (and sometimes untestable) in many empirical studies. We made no assumption about the shape of the relationship between our index and actual true abundance as we cannot know whether the index is driven by changes in local abundance vs. behaviour. Towards the goal of picking an accurate index and validate its use, we highlight the importance of simulations including species interactions and variation in occurrence frequency, in testing different indices of occurrence frequency. Recent advances in using simulations to guide the use of camera trap data (Broadley et al. 2019; Burton et al. 2015; Hofmeester et al. 2017) are promising.

We conducted our study in a large, heavily developed boreal landscape that is typical of areas actively exploited by forestry and oil and gas development. As such, our findings do not necessarily apply in less impacted regions of the boreal forest. In investigating such a large area it is difficult to provide fine-scale sampling resolution; indeed, this is a trade-off with large-scale mensurative studies, and there is no evidence to suggest the resolution of our data affected conclusions. Variables selected for the whole community were the same as variables selected for prey which may be associated with the dominance of prey, especially white-tailed deer, in driving community-level associations in this dataset. The relationship between the area of habitat that best predicts boreal mammal species occurrence is not scaled 1:1 with body size (Fisher et al. 2011), and the signals we detected suggested the data captured at this resolution observed a breadth of processes fitting most species. We plan to, and encourage other research groups to, replicate our study in other boreal landscapes, notably those with different amounts of anthropogenic footprint. Future work in this regard will focus on deploying higher densities of camera-trap stations to better capture differences in home range sizes among species.

Information-theoretic approaches cannot be defined for RDA models because they have no likelihood; hence we used a stepwise selection procedure which, although it did not use AIC, bypassed common issues associated with using significance level as a stopping criterion in the selection. Two well-known risks of using stepwise selection with a stopping significance level are overestimating R^2 (Copas & Long 1991; Diehr & Hoflin 1974; Freedman et al. 1992; Rencher & Pun 1980) and an inflated false positive rate (Derksen & Keselman 1992; Westfall et al. 1998; Whittingham et al. 2006; Wilkinson 1979; Wilkinson & Dallal 1981). However, we mitigated those two limits by following a modified procedure which avoids R^2 overestimation by using a more sophisticated stopping criterion in the forward selection of explanatory variables (two criteria instead of only a significance level), and which avoids Type I error inflation through a systematic test of significance prior to selection, which prevents selection if the test fails (Blanchet et al. 2008). Our preselection of variables to pick the best scale ought not introduce bias or inflate error, because we also used a global significance test for this step, and we used a forward selection which only kept the variable at one scale in the vast majority of cases; we checked collinearity for the very few runs picking several scales. For other models, such as generalized linear models, there are alternatives to using significance level as a stopping criterion in a stepwise selection procedure (or better, two criteria as we did), such as AIC which provides a powerful way to rank candidate models and test hypotheses (Johnson & Omland 2004; Stephens et al. 2005).

Although alternative multivariate modelling approaches that are able to use information-theoretic criteria have recently been developed, they are not expected to outperform RDA in our system. A recent simulation study has shown that a GLM with an AIC-based approach showed similar performance as a RDA with a forward selection approach, in explaining community presence-absence, for low levels of spatial dependence (Carlos-Júnior et al. 2020). As discussed earlier, few MEMs were selected in our study, and their explanatory power was rather low, indicating a generally low level of spatial dependence. If we extend those results to our type of measure, then we would expect a similar or only marginally better result with a GLM with an AIC approach (Carlos-Júnior et al. 2020). Simulations should be conducted to compare these two approaches using a habitat use metric, especially for intermediary levels of spatial dependence where the GLM and AIC approach was better for presence-absence data. The GLM with an AIC approach generally selected fewer variables than truth and the RDA with a forward selection approach selected more variables than truth, despite it using the Blanchet et al. stopping criterion (Carlos-Júnior et al. 2020). Provided we extend the results of that study to our analysis, this could indicate that a fraction of the variables we selected are superfluous. As we did not select many variables, and we were able to understand their selection based on the previous literature, we do not think that selecting superfluous variables was an issue in our case. Overall, the discrepancy in performance and variable selection between both approaches should largely be mitigated by the size of our dataset (Carlos-Júnior et al. 2020; Hegyi & Garamszegi 2011). When all the necessary tools will be available to properly compare both approaches (e.g., tests of the GLM-based method for our type of metric and adaptation of variation partitioning), simulations should be conducted to understand if their performance varies. Then, we would be interested in conducting a short follow-up study comparing results using our dataset and those two approaches. Finally, an important, yet unmitigable, limit of our approach resides in the impossibility to make specific quantitative predictions about other species assemblages (Harris 2015; Warton et al. 2015).

The levels of explained variation that we reached in this study (0.37 for predators, 0.51 for prey, 0.39 for the whole community) are similar to that found in other multispecies mammal studies in the boreal forest, and generally close to that found by similar studies. For example, in a study focusing on the effect of natural and altered habitats on five mammal species in the eastern boreal forest of North America, including four of the species in our dataset, the proportion of explained variation was 0.46 (Bowman et al. 2010). The highest level of explained variance for nine carnivore species, including five of the species in our dataset, in the Canadian Rocky Mountains using both natural and anthropogenic variables was 0.20 (Heim et al. 2019). This study, together with ours, suggests that explaining the landscape use may be harder for predators than for prey, and further justifies our separation in functional groups based on trophic position.

We acknowledge that there may be some sensitivity of our findings to the way we classified groups of species, however there is no strong evidence suggesting that a large part of the unexplained variation in our study is due to modelling choices, such as species grouping. In a similar study, the highest value of explained variation (0.32) was reached for a reduced size-based subset of all predators (Heim et al. 2019), which is substantially higher than for their "global" model (0.20). To evaluate how such a size-based reclassification would impact the level of explanatory power in our study, we did a new analysis where we reclassified each of our functional groups into two groups based on size, and the size of potential prey for predators. One issue for our community is that while both lynx (Squires & Ruggiero 2007; Stephenson et al. 1991) and coyote (Boisjoly et al. 2010b; Crête & Desrosiers 1995; Frenette et al. 2020; Latham et al. 2013b) consume mostly small prey, large prey also figure in their diet, notably in the boreal forest. Therefore, we included them in both our sized-based subgroups. We kept all prey as a predictor for large predators, large predators as a predictor for large prey, small prey as a predictor for small predators, and all predators as a predictor for small prey. The total explained variation was 0.41 for large predators, 0.38 for small predators, 0.56 for large prey, and 0.35 for small prey. Therefore, reclassifying into reduced size-based subsets never outperformed or underperformed our original models by more than 5%.

Beyond species grouping decisions, variation in habitat use or other demographic measures may be generally harder to explain the more species are simultaneously considered, or the more different the considered species are. Indeed, single species GLMs using monthly occurrence to explain species persistence (Fisher & Burton 2018) within the western boreal forest landscape reached an explanatory power (here computed as (null deviance - (model deviance/null deviance)) × 100)) of a maximum of 80% for red fox and a minimum of 29% for fisher with an average of 52% for all mammal species, which although higher, is reasonably close to our own explanatory power.

Generally, at least around half of the variation remained unexplained across studies. We believe there are two non-mutually exclusive explanations for this large amount of unexplained variation, after excluding modelling choices such as species grouping. First, there is evidence that some of the species we studied are ecologically flexible and that, furthermore, the ecological processes we studied impart a substantial amount of variability in species' responses to environment and vary across spatial and temporal scales. This could inherently lower the explanatory power of our models. A second parsimonious explanation is that there is a lot happening in this disturbed boreal forest landscape that we have not been able to measure and use as a predictor in the model. While some studies explicitly included topographical variables (Heim et al. 2019; Wearn et al. 2019), we included it in the creation of our habitat categories (i.e., upland and lowland). However, there are many aspects of the environment we were not able to measure accurately. For example, emerging satellite-based, high-fix-frequency telemetry studies suggest snow and microclimate play a large role in determining animals' site choices on a daily, even hourly, basis. Integrating those variables with the relevant habitat use metric, at the appropriate spatiotemporal scale, is a promising avenue for future research.

The predator-prey system we focused on in this study is likely to react differently in less disturbed landscape for several reasons. As a corollary of the trophic implications of landscape change described earlier, less disturbed landscapes may hold less forage for white-tailed deer and moose. Indeed, many anthropogenic features create space and conditions where early seral vegetation can thrive. In turn, less deer and/or moose is expected to lower the numbers of large predators overall. This would indirectly help sustaining rarer threatened prey affected by apparent competition, like caribou. Less linear disturbance is also expected to decrease the facilitation of predator movement and of caribou habitat fragmentation and functional habitat loss through a decrease in avoidance. Long-term surveys (to account for seasonality and predator-prey demographic cycles) of geographically close areas with differing levels of disturbance are needed to comprehensively understand how disturbance affects predator-prey systems. However, there are a few patterns that have emerged from disturbance studies.

Composition and restoration status of human footprint are key in explaining the effect of disturbance on boreal predator-prey systems. Landscapes with an a priori similar “coarse” amount of disturbance may affect the mammal community very differently. Wolves benefit from trails and seismic lines, while coyotes benefit from roads (Dickie et al. 2016; Fisher & Burton 2018). A landscape with relatively more roads than trails may therefore help coyotes to the potential detriment of wolves. Restored areas are often assumed to provide the benefits of naturally less disturbed areas. Unfortunately, only white-tailed deer reacted negatively to active restoration while wolves and caribou did not react according to expectations, in a landscape close to our study area (Tattersall et al. 2020a). Indeed, although wolf habitat use decreased overall in a landscape-wide active restoration scenario (Beirne et al. 2021), wolves did not avoid restored seismic lines (Tattersall et al. 2020a). Caribou avoided areas with a high density of seismic lines, regardless of restoration status, which does not seem to be a strong predictor for this species (Beirne et al. 2021; Tattersall et al. 2020a). The only prey which consistently favoured restored linear features was possibly the American red squirrel (Beirne et al. 2021), and preference relative to restoration status varied seasonally for many species (Beirne et al. 2021). Explicitly modelling restoration status is especially relevant as vegetation restoration may not always support wildlife conservation in the short term (Hervieux et al. 2014; Serrouya et al. 2021; Tattersall et al. 2020a).

Different levels of disturbance in the boreal forest create novel ecological contexts in which species interactions are altered in complex ways. There is an increasing overlap (i.e., co-occurrence) between wolves and black bears (Leblond et al. 2016) with increasing anthropogenic disturbance, which highlights possible facilitative interactions among large predators (Tattersall et al. 2020b). Coyotes are becoming more abundant than wolves in accessible and southern boreal forest landscapes such as the one we studied (Burgar et al. 2019). However, although coyotes are not yet abundant and competitive with wolves in less disturbed boreal landscapes a few dozen kilometres north of our study area, their ecological flexibility could allow them to still benefit from wolf presence through dynamic and density-dependent behaviour (Beirne et al. 2021; Tattersall et al. 2020b). There are thresholds in predator-prey interactions which mediate the relationship between a species and its environment; this may lead to non-linear relationships (Fisher & Burton 2022, submitted). For example, while there are thresholds in the effect of road disturbance on avoidance by caribou (Leblond et al. 2011), they may be consistently crossed (Leblond et al. 2013).

Our holistic approach involves trading species-specific for community knowledge. Spatial extents have an influence on both the direction and the strength of estimated boreal forest mammal responses to human footprint, suggesting that many relationships are scale-dependent (Toews et al. 2017b). The relationship between space use and the proportion of features at a point in space varies non-linearly across scales and species characteristics and may be relevant to the community at more than one scale (Fisher et al. 2011; Holland et al. 2004). One of the potential limits of our study resides in the fact that the environmental variables in the final models are selected at only one radius, although radii between variables differed a lot (e.g., seven scales for ten variables for predators). This means that selected scales may represent a summary for all species present in the dataset, and that discerning which scale is important for a specific species is less straightforward than when using single-species models. For example, it could be possible that white-tailed deer is associated (positively or negatively) with variable A most strongly at the largest scale and that snowshoe hare is associated with variable A most strongly at the smallest scale. Because collinearity and spatial autocorrelation in habitat and anthropogenic disturbance may not often lead to the selection of the same variable at two different scales, only one scale, possibly intermediate, would be selected.

We believe that looking at different habitat scales around the sampled point is important to address the uncertainty in the scale effects in our model, which cannot be expected to map on directly to home range sizes. We interpreted the relationship between space use and the proportion of habitat types in the surrounding landscape, measured within 20 different surfaces. This is conceptually somewhat different than interpreting home range sizes which are generally better understood as the amount of total area that an individual, or group for gregarious/non-territorial animals, needs to fulfill its life-history requirements. In contrast, the habitat use metric we used summarizes the frequency at which an individual/group uses a point in space which is dependent on movement behaviour and abundance. Such a complex metric will be governed by factors extending well beyond that point in space, for example, whether those individuals/groups have better or worse options elsewhere, whether intraspecific competitors are forcing that choice, whether they are choosing food resources, predation refuge, mates, thermal cover, or some combination of those factors. Because we do not have an accurate understanding about the determining processes shaping the relationship between space use and environmental attributes around sampling sites, the only option we had was to test multiple spatial scales and use the best supported one.

Finally, although some species considered in this study may present individuals or groups with home range sizes larger (e.g., wolf) than the largest surface we considered ($\pi\times5^2$=78.5km^2), or smaller (e.g., squirrel) than the smallest surface we considered ($\pi\times0.25^2$=0.20km^2), we likely have captured a substantial part of the variation in home range sizes and movements among our species with those 20 radii. In other words, the scales we selected for the whole community, could still be meaningful for species with larger or smaller home ranges. For example, we found that roads were important in explaining prey space use at 1.5km (**Table 2-3)** and our result closely match that of previous literature from another part of the range (Leblond et al. 2011).

2.4.5 Applying ecological complexities to managing the boreal forest mammal community

Most of the variation explained by the anthropogenic features submodel is shared with the natural habitat submodel. This, and the fact that natural habitat is the most important prey submodel, could hint that we should generally consider natural habitat as the most limiting factor for prey. Concerning woodland caribou, very large forest patches are needed for populations to thrive, because crowding may facilitate predation and only a little more than half of 100km^2 patches are used by caribou (Lesmerises et al. 2013). Moreover, the minimum amount of undisturbed habitat needed to sustain populations varies a lot according to adult female survival and sex ratio of specific populations (Rudolph et al. 2017). These studies highlighted the need to go beyond a simplistic percent-of-habitat approach and consider the configuration of the landscape. In this chapter, we showed that a variety of anthropogenic features shape the distribution of predators and prey. We acknowledge that anthropogenic features may affect the mammal community at even larger spatial scales than we considered in this study area. The large scale at which some anthropogenic variables were selected for predators, and the ubiquitousness of anthropogenic features, suggest that there might unfortunately be very few areas in our study area without a direct or indirect effect from linear or block features. In other words, a lot of undisturbed habitat may be functionally lost through predator facilitation or perceived risk by caribou. This potentially parallels the results concerning the minimum patch size used by caribou (Lesmerises et al. 2013). Indeed, given the results presented in our study and others, structurally undisturbed habitat patches may need to be very large in order to account for large edge effects. Therefore, conservation efforts aiming to preserve functional undisturbed habitat should systematically consider the indirect effects brought by large-scale negative indirect effects originating from the surrounding matrix. Otherwise, the effective high-quality habitat they protect/restore may be substantially smaller than the actual habitat patch. Another reason to avoid a conservation strategy focusing too narrowly on preserving a species-specific subset of the landscape or a percent-of-habitat approach in undisturbed areas, is that connectivity between habitat patches is key for woodland boreal caribou conservation (Bauduin et al. 2020). Indeed, logging (Courbin et al. 2014), large linear features (Leblond et al. 2011, 2013) and other human infrastructure (Polfus et al. 2011) are known to be avoided by caribou, which could lead to functional habitat loss, through lack of connectivity.

Finally, preserving caribou may not be feasible through protection or restoration of undisturbed habitat, at least in parts of its range. Some authors have argued that addressing the main proximate cause of caribou decline (i.e., predation by wolf) is more successful in the near term than habitat restoration, although the latter must be undertaken too to preserve woodland boreal caribou (Serrouya et al. 2019, 2021). Indeed, predator culls alone are not necessarily successful in helping caribou populations recover (Hervieux et al. 2014), and should only be considered as an unfortunate temporary action designed to "buy time" for other conservation efforts to bring about effective habitat restoration and protection. Combining wolf culls with other mitigation and prevention strategies provided the highest predicted per capita growth rate, which is higher than the threshold representing a 60% probability of a self-sustaining population for several herds in British Columbia (Winder et al. 2020). More heavily disturbed landscapes, if we base the perceived disturbance on caribou growth rates, benefit more strongly from predator culls, but all disturbed landscapes benefit equally when all mitigation measures are considered (Winder et al. 2020). Finally, as stated earlier, restoration is not always successful in hindering predators or attracting caribou back (Tattersall et al. 2020a), and a number of restoration characteristics, such as line sight which reduces caribou habitat use but help deer's (Beirne et al. 2021), must be met for successful management. A recent study conducted on tropical forests showed that while 90% recovery for soil and plant functioning was possible within 25 years, 90% recovery for biomass and species composition took more than a century (Poorter et al. 2021). Although the boreal forest is a very different ecosystem, recovery and restoration take time to be effective, so we need to start efforts as soon as possible.

Both linear and block features influence patterns of the mammal occurrence frequency. Although the effects of block features on species within the boreal mammal community have been researched (Courbin et al. 2009; Courtois et al. 2007; Smith et al. 2000), linear features have been the main focus of recent boreal mammal conservation studies (DeMars & Boutin 2018; Dickie et al. 2016; Latham et al. 2011b; McKenzie et al. 2012; Whittington et al. 2011). Because we found that block and linear features affect predators and prey in different unique ways, mitigation efforts should focus on reclaiming both types of features, to avoid favouring one group of species over another, and to ensure that trophic interaction networks remain intact. While we provide a general overview of how the environment, natural or anthropogenic, influences mammal community variation in human affected landscapes, further opportunities exist to refine this understanding and to improve seasonal and local management recommendations.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Alberta Biodiversity Monitoring Institute, University of Alberta, and Government of Alberta (Environment and Parks, AEP) kindly provided landscape data used in this study. We thank Pierre Legendre and Daniel Borcard for an early review of the methods part of our manuscript. We also thank anonymous reviewers and Martin-Hugues St-Laurent for their very relevant comments, which helped us write a better and more focused manuscript.

Funding sources

Data collection was funded by InnoTech Alberta Inc., AEP, the Petroleum Technology Alliance of Canada AUPRF, MEG Energy and the Alberta Conservation Association. JW was also funded by the Forest Complexity Modelling Collaborative Research and Training Experience Program of the Natural Sciences and Engineering Research Council of Canada. ACB was supported by the Canada Research Chairs program, and JTF by the Oil Sands Monitoring Initiative.

Supplementary material

The following supplementary tables show the results of a species-specific goodness of fit analysis showing the **cumulative** proportion of inertia accounted by species up to chosen axes. 1 would represent a species that is 100% explained by a canonical axis. The analysis was done using the *goodness()* function from the *vegan* package (2.5-2) (Oksanen et al. 2018).

Supp. Table 2-1 : Predator.

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6
black bear	0.407	0.547	0.556	0.560	0.566	0.570
coyote	0.530	0.547	0.580	0.616	0.635	0.639
fisher	0.021	0.030	0.355	0.425	0.455	0.461
lynx	0.252	0.257	0.421	0.470	0.482	0.490
red fox	0.023	0.161	0.162	0.354	0.456	0.457
wolf	0.089	0.389	0.390	0.445	0.447	0.455

Supp. Table 2-2 : Prey.

	RDA1	RDA2	RDA3	RDA4	RDA5
caribou	0.005	0.156	0.267	0.276	0.293
moose	0.001	0.001	0.046	0.128	0.128
red squirrel	0.218	0.254	0.254	0.254	0.278
snowshoe hare	0.669	0.673	0.676	0.676	0.676
white-tailed deer	0.601	0.607	0.618	0.619	0.620

Supp. Table 2-3 : Whole community.

	RDA1	RDA2	RDA3	RDA4	RDA5	RDA6	RDA7	RDA8
black bear	0.022	0.329	0.335	0.336	0.336	0.338	0.338	0.338
caribou	0.012	0.013	0.138	0.201	0.212	0.259	0.282	0.282
coyote	0.112	0.117	0.224	0.226	0.227	0.227	0.233	0.233
fisher	0.029	0.086	0.094	0.119	0.120	0.143	0.149	0.153
lynx	0.217	0.217	0.272	0.279	0.290	0.293	0.294	0.294
moose	0.003	0.015	0.036	0.063	0.081	0.083	0.094	0.097
red fox	0.055	0.076	0.141	0.147	0.235	0.239	0.241	0.241
red squirrel	0.233	0.233	0.251	0.254	0.257	0.266	0.266	0.266
snowshoe hare	0.602	0.602	0.607	0.607	0.608	0.608	0.609	0.609
white-tailed deer	0.636	0.636	0.636	0.636	0.638	0.639	0.640	0.640
wolf	0.000	0.012	0.015	0.135	0.140	0.157	0.157	0.157

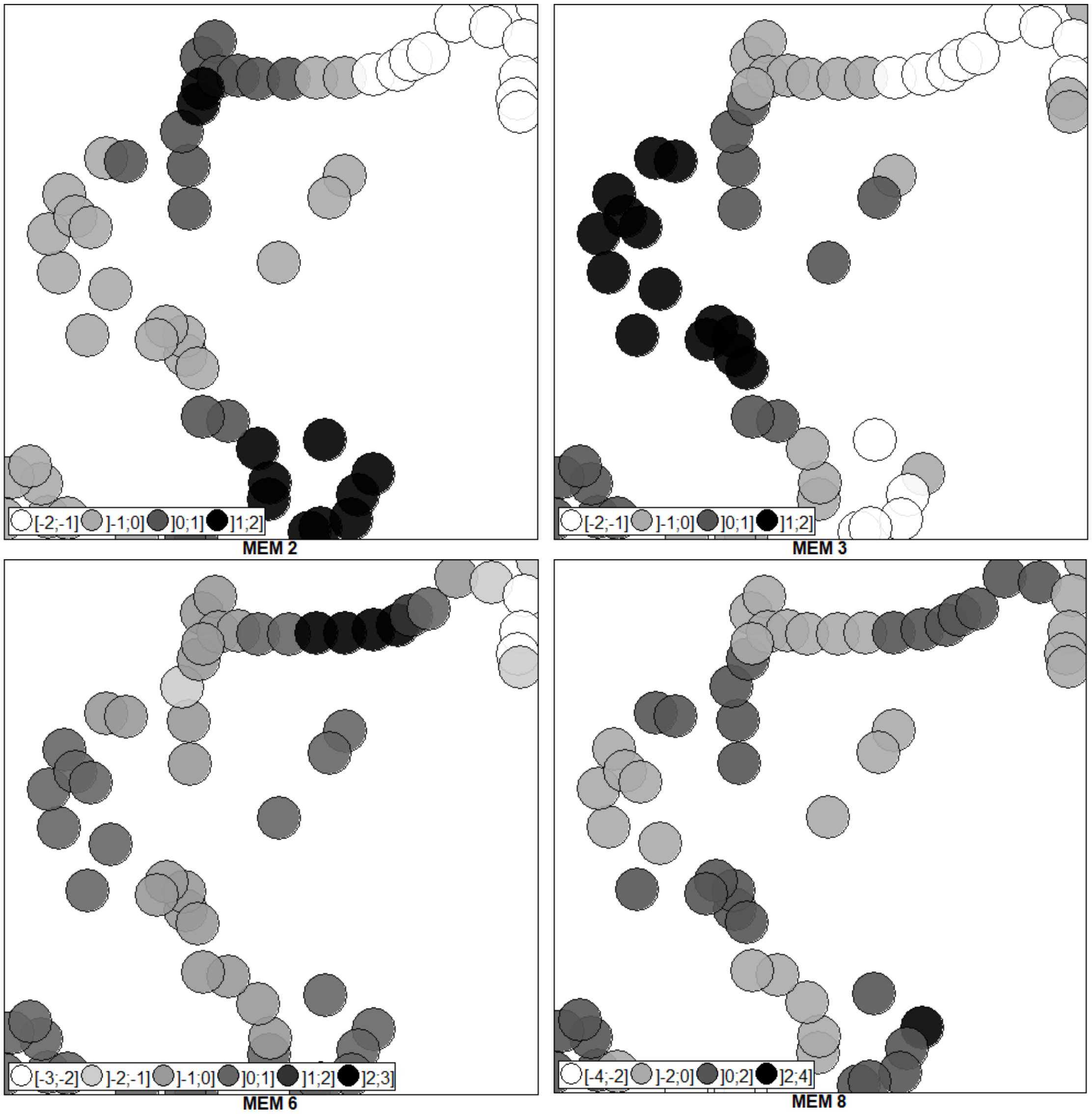
The following are the 1-hour relative abundances values created from the raw data, we used in the analysis:

Supp. Table 2-4 : Transformed relative abundance values.

Cam#	blkbear	coyote	fisher	lynx	rdfox	wolf	caribou	moose	rdsquir	snhare	wtdeer
1	2.00	42.91	0.00	46.86	0.00	4.35	0.00	1.00	7.49	427.73	36.92
2	2.99	4.99	2.17	5.99	0.00	1.32	0.00	0.00	23.99	391.11	52.89
3	91.68	6.00	1.00	4.62	0.00	9.00	3.00	0.00	2.00	32.89	407.08
4	15.00	24.10	5.00	17.01	0.00	4.01	4.00	0.00	4.01	121.05	156.73
5	13.03	8.97	4.01	4.01	0.00	0.00	0.00	0.00	1.51	37.13	164.29
6	8.87	16.27	0.00	8.75	0.00	5.98	0.00	1.49	3.93	2.99	73.85
7	3.99	4.99	0.00	19.96	0.00	2.00	0.00	1.45	3.47	179.63	10.87
8	9.05	3.00	2.10	16.97	0.00	3.08	0.00	2.63	4.00	17.01	111.08
9	2.00	13.02	6.17	18.36	23.99	12.01	1.00	2.00	5.00	34.24	193.13
10	7.01	16.03	4.17	10.81	7.02	7.02	0.00	0.00	10.93	21.79	378.95
11	5.01	4.52	1.00	7.33	1.11	5.47	0.00	2.53	0.00	117.98	244.96
12	20.87	3.93	1.00	2.00	0.00	8.98	0.00	1.00	0.00	14.98	338.99
13	0.00	23.14	2.00	29.35	0.00	3.99	0.00	0.00	3.00	121.97	140.51
14	13.15	8.01	1.00	2.00	0.00	0.00	0.00	0.00	8.01	113.18	100.73
15	1.00	23.02	5.07	8.01	6.01	8.11	0.00	1.00	44.07	268.43	103.31
16	1.00	32.79	1.28	10.03	0.00	10.27	0.00	1.00	1.00	55.77	298.02
17	10.13	11.20	1.36	1.09	1.49	2.99	0.00	0.00	0.00	0.00	372.24
18	0.00	8.97	6.24	1.99	0.00	2.17	10.97	0.00	0.00	23.14	113.77
19	21.97	59.93	0.00	19.92	0.00	3.00	0.00	0.00	0.00	11.01	111.08
20	13.85	43.07	3.00	17.47	0.00	4.00	0.00	0.00	6.00	276.19	234.62
21	1.00	12.02	8.01	3.13	0.00	1.00	3.00	1.35	4.01	57.38	58.95
22	9.02	16.04	4.52	18.06	0.00	2.19	1.00	1.47	1.11	25.46	353.95
23	7.65	9.98	2.62	3.99	0.00	5.98	1.49	0.00	1.00	116.65	179.71
24	2.99	22.95	1.16	28.94	0.00	8.39	4.35	1.09	5.65	117.87	76.98
25	11.98	170.26	1.00	19.01	0.00	38.03	0.00	0.00	25.02	203.32	324.52
26	52.03	18.00	1.00	1.27	1.00	6.26	0.00	1.00	0.00	15.02	140.22
27	1.35	13.03	5.47	11.53	0.00	3.01	0.00	0.00	1.00	18.05	200.46

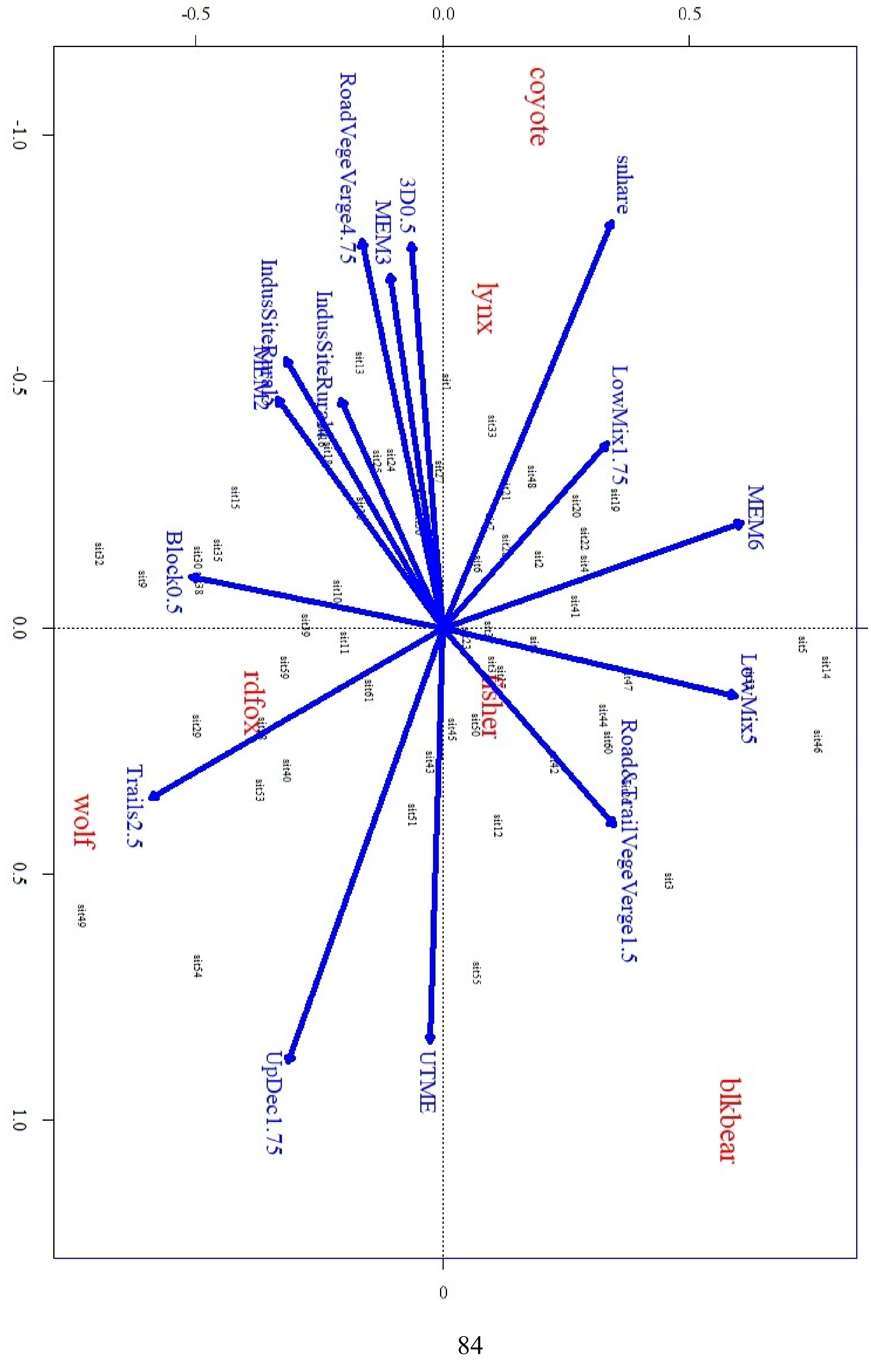
28	8.80	24.40	5.09	10.85	0.00	5.99	0.00	0.00	5.96	84.83	302.41
29	6.99	2.99	2.99	8.74	2.00	16.97	0.00	1.00	0.00	37.92	431.59
30	2.17	15.83	3.00	13.00	0.00	24.02	0.00	1.05	5.26	109.14	118.08
31	24.02	7.00	2.00	35.06	3.06	6.34	6.16	0.00	1.00	26.99	142.10
32	0.00	2.00	2.00	1.00	0.00	3.84	2.00	0.00	0.00	30.08	125.70
33	6.02	37.09	0.00	57.27	0.00	4.07	0.00	11.04	3.80	39.93	67.21
34	6.56	10.44	4.99	15.71	0.00	7.49	0.00	2.00	2.99	44.87	222.34
35	5.98	33.34	0.00	16.20	0.00	37.99	0.00	2.99	8.99	36.19	118.86
36	6.29	40.78	0.00	7.00	0.00	15.02	0.00	1.00	0.00	85.06	231.37
37	16.45	12.00	5.00	23.02	0.00	10.02	0.00	3.80	1.00	198.32	155.25
38	4.17	18.03	1.00	10.93	0.00	28.64	0.00	2.56	5.01	50.13	43.11
39	4.01	7.60	2.22	5.09	0.00	10.47	0.00	1.36	0.00	34.29	253.48
40	10.98	6.98	6.98	1.00	4.99	11.57	0.00	1.25	1.00	1.45	542.89
41	9.98	11.99	1.13	5.99	0.00	3.00	0.00	3.00	1.00	71.03	227.01
42	22.02	7.69	6.00	4.01	2.06	4.00	0.00	0.00	3.06	94.98	366.92
43	11.41	2.00	2.00	7.01	0.00	8.20	0.00	3.00	1.35	24.04	390.98
44	17.94	9.02	7.02	5.01	0.00	4.07	0.00	1.51	1.47	25.34	357.12
45	21.70	10.94	8.94	7.98	0.00	15.95	0.00	2.62	1.00	38.88	416.76
46	13.73	3.99	5.80	2.00	0.00	0.00	0.00	0.00	0.00	20.98	194.50
47	12.00	11.98	9.21	1.46	0.00	2.00	0.00	1.00	0.00	13.85	167.11
48	10.02	48.97	1.00	14.00	0.00	5.01	0.00	1.00	0.00	171.12	157.25
49	10.02	0.00	5.01	1.00	7.69	21.05	0.00	0.00	0.00	1.00	366.92
50	55.77	14.67	6.33	39.93	0.00	30.94	0.00	0.00	2.19	69.99	238.53
51	43.20	5.99	3.99	10.97	11.23	14.97	0.00	1.99	1.00	7.98	185.52
52	12.73	1.09	1.00	15.83	0.00	0.00	0.00	0.00	5.24	121.75	38.15
53	10.01	3.00	6.16	3.00	0.00	21.00	0.00	1.00	3.08	79.58	152.97
54	4.00	0.00	1.00	0.00	0.00	7.02	0.00	9.01	0.00	2.70	144.23
55	24.59	0.00	1.00	1.00	0.00	8.87	0.00	13.03	1.00	13.20	186.21
56	4.07	16.40	4.37	13.42	0.00	6.98	0.00	0.00	1.00	41.91	237.30
57	5.98	20.94	0.00	18.84	0.00	7.92	0.00	1.00	0.00	4.35	683.22
58	9.99	2.10	6.99	15.79	9.01	11.01	0.00	1.46	6.00	7.00	809.50
59	7.01	3.08	0.00	14.00	0.00	14.02	0.00	2.00	0.00	5.00	327.22
60	35.06	8.35	6.76	15.02	0.00	7.02	12.02	1.00	78.18	135.34	54.14
61	5.01	5.02	4.40	2.53	0.00	7.45	0.00	4.44	0.00	13.12	89.67

The following is a simple representation of what MEMs represent in our study landscape with MEM values at each of other sites. MEM 1 captures spatial autocorrelation at a broad scale, and as you can see the variation in the spatial process it captures (represented by shades of grey; from white to black) is "gradual" over the whole landscape, from highest value in the north-western sites, to lowest in the south-eastern sites. The following MEMs capture spatial autocorrelation at finer scales. Please refer to (Legendre & Legendre 2012) or to (Dray et al. 2006) for a proper introduction.



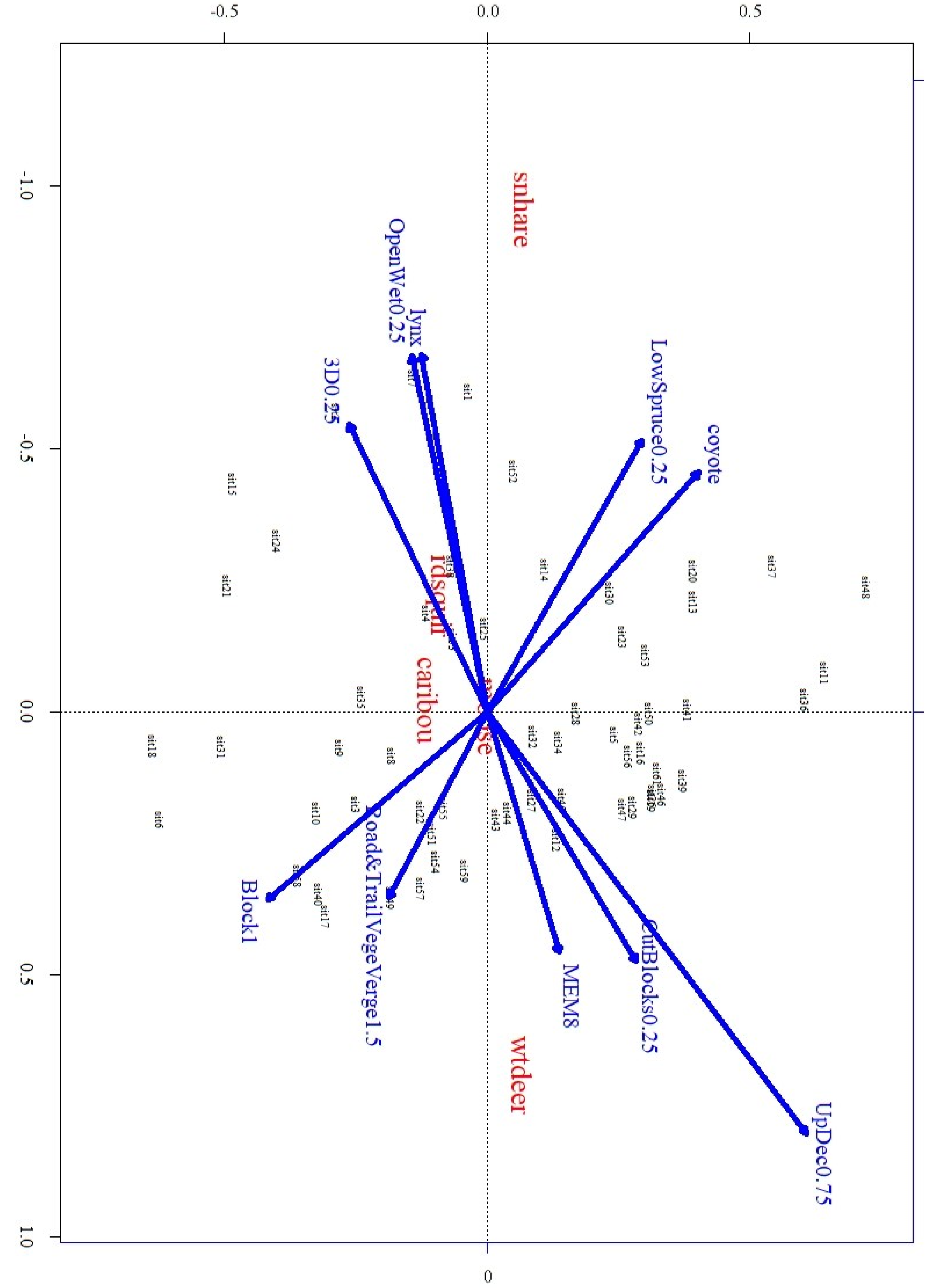
Supp. Figure 2-1 : Moran's Eigenvector Maps 2, 3, 6 and 8 for our study area.

A)

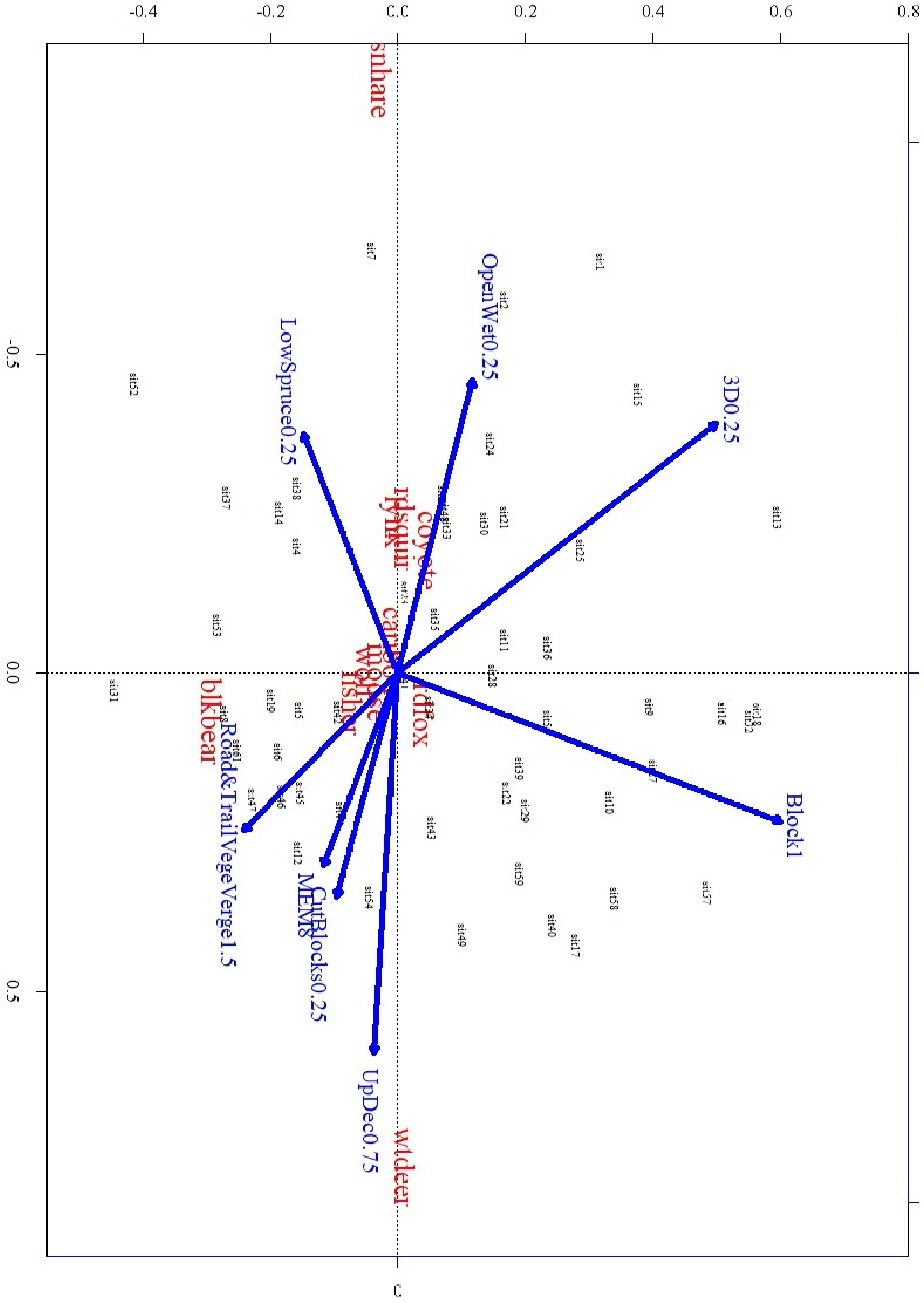


-0.5
0.0
0.5
-1.0
-0.5
0.0
0.5
1.0
0
coyote
lynx
fisher
rdfox
wolf
blkbear
snhare
RoadVegeVerge4.75
3D0.5
MEM3
IndusSiteRural
IndusSiteRural1
MEM2
LowMix1.75
MEM6
Block0.5
LowMix5
Road&TrailVegeVerge1.5
Trails2.5
UpDec1.75
UTME
sit1
sit13
sit33
sit48
sit19
sit20
sit22
sit24
sit25
sit27
sit21
sit6
sit2
sit15
sit35
sit32
sit30
sit9
sit10
sit41
sit39
sit11
sit59
sit61
sit5
sit14
sit47
sit44
sit60
sit46
sit29
sit45
sit50
sit43
sit40
sit53
sit42
sit51
sit12
sit3
sit49
sit54
sit55

B)


snhare
lynx
OpenWet0.25
3D0.25
LowSpruce0.25
coyote
rdsquir
caribou
moose
wtdeer
Block1
Road&TrailVegeVerge1.5
MEM8
CutBlocks0.25
UpDec0.75
-0.5
0.0
0.5
-1.0
-0.5
0.0
0.5
1.0
0

C)



-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
-0.5
0.0
0.5
0
snhare
coyote
rdsquir
lynx
fisher
redfox
blkbear
wtdeer
OpenWet0.25
LowSpruce0.25
3D0.25
Block1
Road&TrailVegeVerge1.5
CutBlocks0.25
MEM8
UpDec0.75

Supp. Figure 2-2 : Triplots of the relationship between habitat use and environmental variables for each analysis (A) for predators; B) for prey; C) for the whole community). Species names are indicated in red. Environmental variables are indicated in blue. Sites are indicated in black. The arrow length and direction of the variables correspond to the variance that can be explained. The angle between an explanatory variable x and a response variable y in the biplot reflects their correlation. That is, the smaller the angle, the stronger the correlation. We did not plot red arrows for species in our triplots for easier readability, but they would be proportional to the position of species labels. Please refer to (Legendre & Legendre 2012) for an in-depth explanation of RDA triplots.

3. CHAPITRE 3 – Modélisation de la connectivité génétique du paysage du dendroctone du pin ponderosa dans l’ouest du Canada.

Modelling landscape genetic connectivity of the mountain pine beetle in western Canada

Julian Wittische[1*], Jasmine K. Janes[2,3], Patrick M. A. James[1]

[1] - Department of Biological Sciences, Pavillon Marie-Victorin, Université de Montréal, 90 Vincent-d’Indy Ave, Montréal, QC H2V 2S9, CANADA

[2] - Vancouver Island University, Nanaimo, BC. CANADA

[3] - School of Environmental and Rural Sciences, University of New England, Armidale, NSW. AUSTRALIA. 2350.

Published (invited to the MPB system special issue)

Julian Wittische, Jasmine K. Janes, and Patrick M.A. James. 2019. Modelling landscape genetic connectivity of the mountain pine beetle in western Canada. Canadian Journal of Forest Research. 49(11): 1339-1348. https://doi.org/10.1139/cjfr-2018-0417

Abstract

The current mountain pine beetle (MPB; *Dendroctonus ponderosae* Hopkins, 1902) outbreak has reached more than 25 million hectares of forests in North America, affecting pine species throughout the region and substantially changing landscapes. However, landscape features that enhance or limit dispersal during the geographic expansion associated with the outbreak are poorly understood. One of the obstacles in evaluating the effects of landscape features on dispersal is the parameterization of resistance surfaces. This parameterization often relies on expert opinion. Unfortunately, the incompleteness of expert knowledge about species that are difficult to monitor could lead to erroneous guidance about landscape connectivity for natural resources managers, which crucially needs to be avoided given the current biodiversity crisis. Indeed, researchers often need to make a priori assumptions about complex, non-linear, non-monotonic movement behaviour in novel habitats when building landscape connectivity models. In this study, we assessed the impact of four environmental variables on MPB genetic connectivity across western Canada. We optimized resistance surfaces using genetic algorithms and models of maximum likelihood population effects, based on pairwise genetic distances and ecological distances calculated using random-walk commute-time distances. Unlike other methods for the development of resistance surfaces, this approach does not make a priori assumptions about the direction or shape of the relationships between environmental features and their cost to movement. We found highest support for a composite resistance surface including elevation and climate. These results further the understanding of MPB movement during an outbreak. Additionally, we demonstrated how to use our results for management purposes.

Keywords: Range expansion; Gene flow; Random-walk resistance; Linear mixed-effect model; Insect outbreaks

3.1 Introduction

Dispersal is an important determinant of ecological and evolutionary dynamics due to its influence on population connectivity (Taylor et al. 1993). In turn, connectivity has significant implications for population (Martin & Fahrig 2016) and species persistence (Thomas 2000). Understanding dispersal is also important for the effective management and conservation of populations and communities faced with ongoing global change and increasingly fragmented and degraded habitats (Haddad et al. 2015). Dispersal models are particularly needed to help us better forecast range expansions of alien invasive and native irruptive species (Rejmánek & Richardson 1996). Irruptive and invasive species represent non-negligible threats to biodiversity and the provisioning of ecosystem services at a global scale (Simberloff et al. 2013). Improved understanding of how abiotic and biotic conditions influence dispersal of outbreaking or invasive species is of fundamental value to natural resource managers.

One species of particular concern in the boreal forest ecosystem of western North America is the MPB (*Dendroctonus ponderosae* Hopkins, 1902; ITIS.org Taxonomic Serial Number 114918). The MPB is a highly mobile, native, and irruptive forest insect pest whose outbreaks have significant ecological and economic consequences as it feeds on the majority of pine species in its range, including, among others, lodgepole pine (*Pinus contorta* Douglas ex Loudon), sugar pine (*Pinus lambertiana* Douglas), western white pine (*Pinus monticola* Douglas ex D. Don), and ponderosa pine (*Pinus ponderosa* Douglas ex P. Lawson & C. Lawson), and is able to maintain outbreaks in healthy stands (Safranyik & Carroll 2006). MPB usually attack weak or damaged trees as do other bark beetle species; however, once MPB populations increase beyond a threshold, MPB populations are able to overcome tree defences. MPB populations may then erupt into self-propagating large-scale outbreaks (Safranyik & Carroll 2006). MPB outbreaks represent one of the most significant biotic threats to Canadian forests (Boucher et al. 2018; Raffa et al. 2008). Since the early 2000s, outbreaking populations of the MPB have undergone rapid expansion from their historical range (Cullingham et al. 2011; de la Giroday et al. 2012). Eastward outbreak spread rates exceeded 80 km·year^{-1} in certain years (Cooke & Carroll 2017).

Models combining atmospheric dispersal and weather found that long-distance dispersal events greater than 50 km are plausible (Ainslie & Jackson 2011). At the new edge of its range, close to the Alberta-Saskatchewan border of Canada, the outbreak has now reached novel habitats and has the potential to spread further in the boreal forests, both eastward and northward, and is currently threatening forest resources in these areas (Cullingham et al. 2011; Janes et al. 2014; Safranyik et al. 2010a; Sambaraju et al. 2012).

Despite being able to monitor the extent and rate of spread of the current outbreak (e.g., Cooke & Carroll 2017), we have limited knowledge about what hinders or facilitates MPB dispersal and movement and, thus, population connectivity (Taylor et al. 1993). Previous studies conducted in western Canada have shown that MPB populations belong to two clusters, a northern cluster and a southern one (James et al. 2011a; Janes et al. 2014; Samarasekera et al. 2012; Trevoy et al. 2018). The earliest study compared two geographic subregions based on the progression of the outbreak from historically documented areas (southern region) to areas where the beetle had not previously been recorded (northern region). This progression was described in the literature based on the presence of beetles and damage to trees, notably from aerial surveys (Aukema et al. 2006; Raffa et al. 2008; Safranyik et al. 2010b) as well as based on molecular evidence (Mock et al. 2007). Their analysis of genetic data supported their partition based on natural history, as it showed that the variation between the northern beetles and the southern beetles was significant. Furthermore, James et al. 2011 found the same pattern of variation between regions for a fungal symbiont of the MPB, which overall suggested the existence of two MPB genetic clusters. The following study (Samarasekera et al. 2012) went further by empirically identifying the previously assumed genetic clusters and their geographical boundary. Sampling a larger geographical area than the previous study, they found support for the existence of two clusters with a boundary only slightly north of that which was chosen in the previous paper. The northern cluster showed reduced genetic diversity relative to the southern one, which indicates a more recent colonization (Samarasekera et al. 2012). Two recent studies (Janes et al. 2014; Trevoy et al. 2018) supported the existence of the two genetic clusters described earlier, while providing additional information about adaptive genetic variation and the origin of beetles attacking a national park. In addition to all those studies focusing on the outbreak in western Canada, two continent-wide (including US populations) studies also confirmed the split between the northern and southern clusters (Batista et al. 2016; Janes et al. 2018).

The importance of environmental features such as elevation, climate, and host volume on MPB dynamics and local population connectivity has been described previously (Bentz et al. 2010; de la Giroday et al. 2011; James et al. 2011a). Drought has also been shown to increase host tree vulnerability and may influence MPB population connectivity (Berg et al. 2006; Raffa et al. 2008; Safranyik et al. 1975). Although multiple landscape features likely influence MPB population connectivity (**Table 3-1**), we have yet to quantify the relative importance of each of these features to MPB movement across the Canadian portion of its range and how they influence outbreak spread.

Table 3-1: Summary of predictor variables used in the optimization and selection process and the associated hypotheses.

Data	Sources	Hypothesis	References
Elevation	SRTM, GMTED NRCAN registry.opendata.aws open.canada.ca	(-): limited dispersal at high elevation. (-): relief could act as a barrier for wind-dispersed MPB.	(Amman et al. 1973) (de la Giroday et al. 2011)
Pine volume	NRCAN/CFS open.canada.ca	(+): preference for high-volume stands could increase effective dispersal.	(James et al. 2011a; Safranyik et al. 2010a)
Drought	NRCAN/CFS open.canada.ca	(-): severe drought decreases survival and brood production by depleting or drying phloem tissue (+): drought weakens trees which lowers their defenses and could increase effective dispersal.	(Amman 1972; Safranyik & Carroll 2006) (Berg et al. 2006; Raffa et al. 2008)
Climate suitability index	NRCAN/CFS	(+): climatic limitations to successful reproduction would hinder effective dispersal.	(Bentz et al. 2010; James et al. 2011a; Shore et al. 2004)

SRTM: Shuttle Radar Topography Mission (NASA and NGA)
GMTED: Global Multi-resolution Terrain Elevation Data (USGS and NGA)
NRCAN: Natural Resources Canada
CFS: Canadian Forest Service

Due to the challenges of directly monitoring small organisms such as forest insects, population connectivity is increasingly estimated using indirect methods that quantify intergenerational gene flow using molecular markers (Broquet & Petit 2009). Dispersal can be considered effective when it creates gene flow, which for mobile animal species often requires dispersing individuals to reproduce in a population after having dispersed (Broquet & Petit 2009). Because the genetic information used by indirect methods that estimate dispersal is affected by gene flow, they provide estimates of "effective" dispersal (Clobert et al. 2009). In general, the greater the gene flow is, the greater the presumed effective dispersal and, hence the greater the functional connectivity between populations (Clobert et al. 2009).

When gene flow is combined with information on landscape resistance within a landscape genetics framework (Manel et al. 2003), one can also infer which landscape features hinder or facilitate movement among populations. Resistance may represent the physiological cost of moving through the landscape, the mortality for the organism moving through the landscape, and (or) the behavioural response of an organism to crossing the landscape (Bonte et al. 2012).

In this study, we investigate how landscape resistance influences MPB population genetic connectivity using landscape genetics to increase our understanding of the spatial dynamics of MPB outbreaks. We specifically sought to identify the landscape features that have affected MPB movement in the recent outbreak region in western Canada.

An important challenge in resistance-based models of landscape connectivity is that model outcomes are sensitive to the parametrization of resistance surfaces (Rayfield et al. 2010; Spear et al. 2010). Often, such parameterization relies heavily on expert opinion, which has been shown to be unreliable at times (Koen et al. 2012; Rayfield et al. 2010) and occasionally to perform worse than random null models of resistance (Charney 2012). To overcome the issues associated with parameterizing of resistance surfaces, we used a machine-learning approach (Peterman 2018) to identify the optimal model of MPB landscape genetic connectivity. We compared millions of potential resistance surfaces resulting in a link-based landscape genetics model of MPB genetic connectivity. We sought to identify the best fit between genetic and ecological distance matrices. The identified best resistance surfaces were used to predict future expected population connectivity and likely routes of expansion. This information could allow for a better prioritization of preventive management efforts.

3.2 Methods

3.2.1 Study area and genetic data

To address our research questions, we used the genetic dataset of Janes et al. (2014). These data include 532 MPB individuals from 27 sites throughout Alberta and British Columbia, Canada (**Figure 3-1**, **Table 3-2**). These sites included both historical and recently colonized parts of the MPB range circa 2006-2010 (Janes et al. 2014). We used the same 764 single nucleotide polymorphisms (SNPs) that had been manually scored and screened for linkage disequilibrium and deviation from Hardy-Weinberg equilibrium (Janes et al. 2014). Allele frequencies for all populations were used to estimate genetic differentiation between populations using a sample size corrected estimator of F_{ST} (Weir & Cockerham 1984). The number of individuals and observed heterozygosity are displayed in **Table 3-2**. Additional details can be found in Janes et al. (2014).

3.2.1 Environmental data

We investigated the influence of elevation, climate, drought, and pine volume on population connectivity quantified using a matrix of pairwise F_{ST} values (**Table 3-1**). Environmental data were represented as continuous raster layers for the full extent of the study area at a spatial resolution of 10 km. We chose an elevation raster built from multiple sources (**Table 3-1**) and retained areas above sea level. Pine volume represents the volume of pine trees (m^3) per hectare. We used the climate moisture index to represent drought (Hogg 1997). This index measures the absolute difference between precipitation and potential evapotranspiration. Finally, we used the mean Climate Suitability Index (CSI) to incorporate the effect of climate on the MPB.

Table 3-2 : Sampling size and observed heterozygosity for all sampling sites.

Sampling site	Sample size	Observed heterozygosity	Longitude	Latitude	Year
Canmore (CAN)	9	0.371	- 115.3364	50.9323	2010
Crowsnest Pass (CPS)	21	0.355	- 114.5525	49.6574	2007/2008
Cypress Hills (CYH)	18	0.357	- 110.0363	49.5931	2007
Fairview (FAV)	21	0.364	- 119.3860	56.5994	2008
Fox Creek (FOX)	23	0.362	-116.6348	54.4806	2008/2010
Kakwa-Wilmore (KAW)	21	0.376	- 119.6004	53.8036	2006/2008
Grande Prairie (GRP)	21	0.364	- 118.6135	54.9924	2008/2010
Cranbrook (CRA)	20	0.362	- 115.6460	49.4086	2010
Ft. St. James (FSJ)	20	0.355	- 121.7120	56.7043	2006
Ft. St. John (FTJ)	19	0.363	- 124.4203	54.6452	2006
Golden (GOL)	21	0.370	- 116.3816	51.0744	2007
Houston (HOT)	21	0.360	- 126.6527	53.9940	2006
Kelowna-Peachlands (KPE)	21	0.370	- 119.6690	49.9965	2006/2010
Kootnay-Yoho (KOY)	20	0.365	- 116.2908	51.1229	2006/2007
Lac Le Hache (LAC)	20	0.389	- 121.5984	51.7307	2006
Manning Park (MAP)	21	0.353	- 121.0697	49.2162	2006
McBride (MCB)	19	0.379	- 120.1266	53.3116	2006
Prince George (PGE)	17	0.377	- 122.8080	53.9065	2006
Quesnel (QUE)	20	0.385	- 122.2741	53.0370	2006
Smithers (SMI)	21	0.356	- 127.3505	54.9289	2010
Tatla Lake (TAT)	21	0.372	- 124.4130	51.9715	2006
Terrace (TER)	16	0.363	- 128.5000	54.8365	2010
Tumbler Ridge (TUR)	21	0.355	- 121.9848	55.5387	2010
Valhalla (VAL)	18	0.370	- 117.5181	49.7503	2006
Valemount (VMT)	22	0.383	- 119.3816	52.8532	2007/2010
Wells Grey (WEG)	20	0.386	- 120.0120	51.7411	2006
Whistler (WHI)	20	0.338	- 122.9251	50.1678	2006

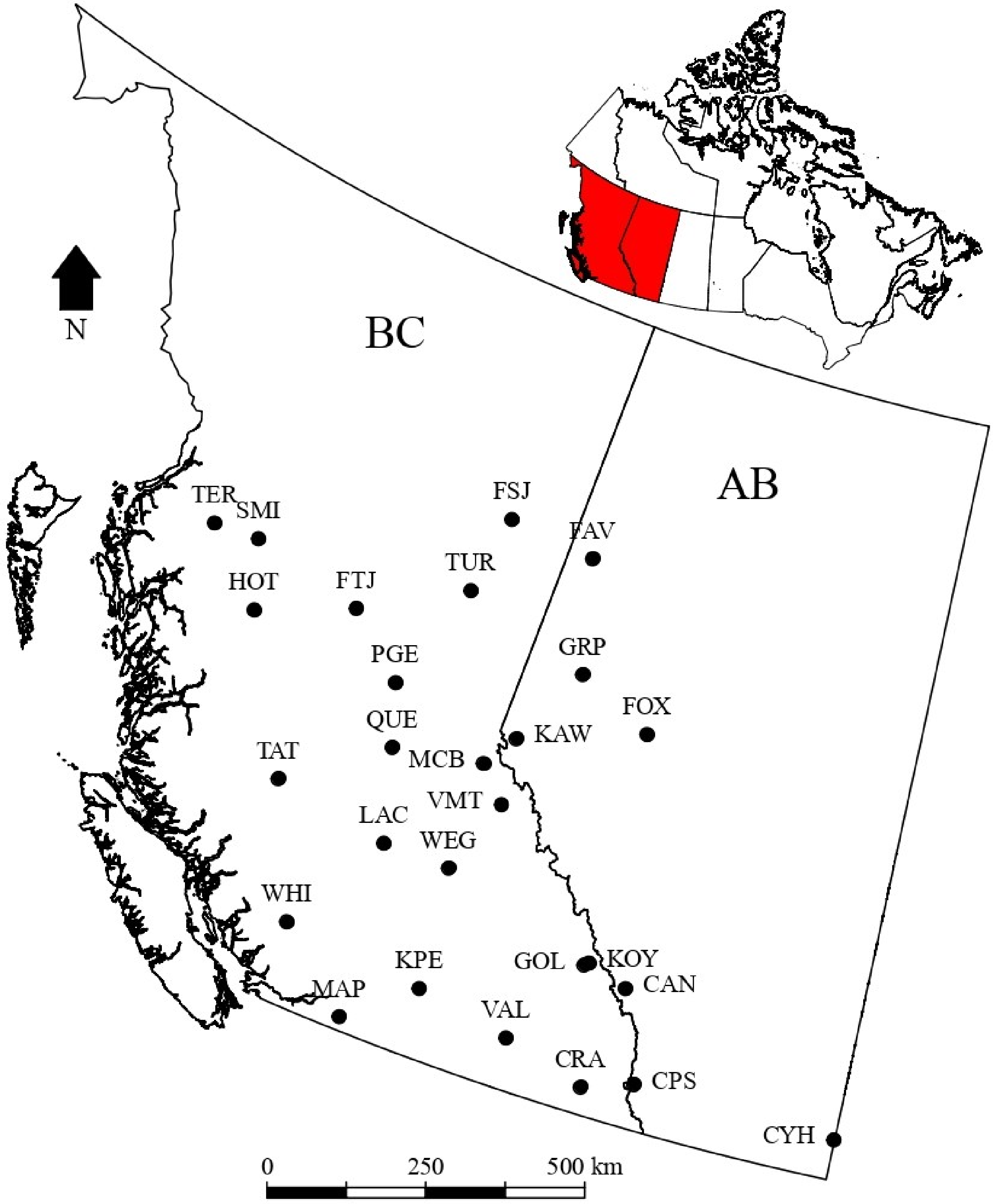


Figure 3-1: Populations sampled across Alberta and British Columbia. Sampled sites are represented by black circles. See **Table 3-2** for site abbreviations.

CSI is a synthetic measure built using multiple climatic variables known to determine MPB development, survival, and attack success. Three models have been developed to evaluate the climate suitability of the boreal forest for mountain pine beetle outbreaks (Safranyik et al. 2010b). We chose the CSI model (Carroll et al. 2004) because, although all three models may contribute to better understanding of the MPB potential for range expansion (Safranyik et al. 2010b), it is the most comprehensive. Indeed, CSI integrates a diversity of climatic variables which influence key demographic processes in the life cycle of the beetle and does not focus on one stage or one process as the other models (Carroll et al. 2004; Safranyik et al. 1975). Specifically, CSI is based on six climatic variables (Carroll et al. 2004). They notably take into account that temperatures inferior to -40°C under the bark are completely lethal (Safranyik & Llnton 1998), that enough degree-days are needed to lead to a temporally stable univoltinism (Logan & Powell 2001), which is necessary for outbreaks, and that late summer temperature must be high enough for flight (McCambridge 1971), which is required for successful mass attacks and, therefore, spread. CSI also includes the fact that water availability affects the beetle, its host, and its fungi symbiont, and that consecutive years presenting below-average precipitation are associated with demographic increases (Thomson & Shrimpton 1984).We used the CSI mean of the period from 1991 to 2020, to cover both the start of the current outbreak and the start of the expansion to Saskatchewan.

We made no a priori assumptions regarding the magnitude, shape, scale, or direction of influence of environmental resistance on gene flow. Instead, through application of our chosen optimization algorithm, we let the data tell us how environment affects MPB gene flow.

3.2.2 Optimization process

We used a genetic algorithm optimization approach provided by the *ResistanceGA* 3.0.0 R package (Peterman 2018) to convert information on landscape spatial environmental heterogeneity into movement costs (Spear et al. 2010; Zeller et al. 2012). Genetic algorithms are an example of machine learning that represents a suite of general approaches used to extract functional relationships from data without prior assumptions (Hastie et al. 2017). Machine learning methods have been incorporated in the biologist toolset with very diverse applications (Peters et al. 2014; Tarca et al. 2007). For example, in this study, we used machine learning to improve resistance surfaces in an iterative way. The genetic algorithm that we used is an evolutionary example of machine learning algorithm. It is evolutionary in that we can consider the different resistance surfaces, as individuals in a population undergoing evolution. The evolution starts with our population (i.e., set of resistance surfaces) which is part of a generation. As in a natural population, the fitness (model performance) of individuals varies. The fittest individuals survive natural selection (model performance threshold) and the characteristics (parameters) that allowed them to survive are therefore preserved. The genomes of surviving individuals are then mutated and recombined: a new generation has been created. Mutation represents a divergent force which partly changes the parameters of the top resistance surfaces to avoid local parameter optima and explore a different parameter space. Crossover represents a convergent force that will concentrate the characteristics of two good resistance surfaces to produce possibly superior new solutions. This whole process is repeated for a number of generations until the population reaches a fitness level that does not improve for several generations. The *ResistanceGA* 3.0.0 R package (Peterman 2018) that we used depends on the *GA* package (Scrucca 2013, 2017) to implement its genetic algorithms.

Throughout the optimization process, genetic distances were regressed against ecological distances using linear mixed-effects models with a maximum likelihood population effects parameterization, an approach that overcomes the issue of non-independence of pairwise distances (MLPE; Clarke et al. 2002, Row et al. 2017, Shirk et al. 2018). MLPE does so by including a population covariance random effect term that accounts for the non-independent error structure associated with pairwise distances (Clarke et al. 2002).

We calculated pairwise matrices of ecological distances between sample sites through a random-walk commute-time algorithm. Commute-time distances represent the expected length of paths travelled by random walkers during a round trip between two nodes (Göbel & Jagers 1974) and are proportional to resistance distances calculated through electrical circuit theory (Doyle & Snell 1984). Although several environmental variables can contribute to the surface used in a model (composite surface), one predictor is used. In addition to all the resistance surfaces built from landscape features, we examined an intercept-only model (null model), as well as a simple geographical distance surface where the resistance of all cells in the resistance surface is set to one (i.e., isolation-by-distance). This gave us a total of 17 different resistance models to parameterize through the genetic algorithm. Single-surface models were used to evaluate the individual hypotheses described in **Table 3-1**. Following individual optimizations, we constructed and evaluated multiple composite surfaces (i.e., surfaces made up of different combinations of the individual surfaces) to identify an overall best integrated model.

The first step of the optimization process for a single surface is to generate a random initial set of resistance surfaces. These surfaces are created by applying a transformation to each spatial environmental variable that is hypothesized to influence genetic connectivity. Possible transformation functions included eight exponential-based functions, each of which is defined by two parameters: shape and maximum resistance. We used saturating monotonic functions (origin-fixing linear left end, a saturating middle, and an asymptote right end) and unimodal functions (origin-fixing linear left end, a hump-shaped middle, and a zero-approaching right end), as well as the reverse, inverse, and inverse-reverse of both functions (Peterman 2018). For each transformation, initial values for the shape and maximum resistance parameters were chosen randomly (*ResistanceGA* defaults; Peterman 2018). At the end of this first step, we evaluated the set of potential resistance surfaces for their ability to model our pairwise genetic response matrix on the basis of their log-likelihood values. The top 5% of those resistance surfaces were retained (*ResistanceGA* defaults; Peterman 2018).

Next, we sought to identify global parameter optima for this set of retained resistance surfaces using genetic algorithms. Optima were identified through “evolution” of the model parameters through the processes of “mutation” (probability = 0.2) and “crossover” (probability = 0.9). All steps are repeated until no improvement in log-likelihood was found for 25 iterations. This process was applied to each landscape variable and combinations of variables, giving us 17 parameterized candidate models.

3.2.3 Model selection and performance

We selected the best model from our set of candidate models using the Akaike information criterion corrected for sample size (AICc) and associated Akaike weights (ω_{AICc}). A resampling analysis was then conducted to validate the selection of our models (Peterman 2018). The goal of the resampling procedure was to assess how sensitive our conclusions were to outliers (sites). To do this, 75% (*ResistanceGA* defaults; Peterman 2018) of our populations were randomly resampled without replacement 1000 times. For each iteration (i.e., sample) of the resampling, we ranked the previously optimized models according to their log-likelihood when using this sample (1 meaning the best model and 2 meaning the second-best model), and recorded which model was the best model. Average ranking and frequency of model being the top model were calculated for each model over all 1000 resampling iterations. To assess the absolute performance of the top model, we evaluated its marginal R^2 (fixed factors) and its conditional R^2 (fixed and random factors).

Finally, we sought to verify whether our best model based on all populations (**Figure 3-1**), would still be considered a top model when examining population connectivity only within the northern genetic cluster. This cluster is of special interest because the current MPB outbreak is expanding from this cluster (James et al. 2011a; Janes et al. 2014; Samarasekera et al. 2012). To do so, we reran the resampling analysis described above only within this cluster.

3.2.4 Model applications – Origin of the Hinton MPB population and Potential for eastward expansion

For both applications, we modelled connectivity using circuit theory and visualized our previous results using electrical current maps (McRae et al. 2008). Here, electrical current is analogous to the probability of movement through the landscape (Doyle & Snell 1984). The resulting current map illustrates the most likely paths of colonization (McRae et al. 2008). For both applications, we used infested areas as sources of current and areas at future risk of being attacked as grounds in virtual electrical circuits connected by resistors based on the values of the best model obtained.

Forested areas in and around the municipality of Hinton, Alberta, which depends heavily on forestry, are currently at risk of being attacked by the MPB (Weber 2017). Local government has great interest in understanding from where an MPB outbreak would most likely originate. We therefore sought to demonstrate the utility of our final model of MPB connectivity to predict relative genetic connectivity between Hinton and several potential source populations: Jasper, Edson, and Grande Prairie (**Figure 3-4**). Based on the current maps, we then infer the likely origin of beetles moving into the Hinton area.

We also applied our approach to forecast MPB landscape connectivity between the outbreak range in Alberta in 2008 and forested areas east of there that represent potential further range expansion (Safranyik et al. 2010a). For this larger scale application, we set the area damaged by the MPB in 2008 (**Figure 3-5**) as an electrical source and the Saskatchewan/Manitoba border as a ground in an electrical circuit. Movement paths were only considered possible over parts of the landscape that contained pine. Because of the possibility of pine volume becoming more important to MPB spread in areas east of the Rockies (Cooke & Carroll 2017; Safranyik et al. 2010a), we forecast future MPB spread using a pine-only optimized resistance surface, in addition to the best composite surface obtained through the model selection.

3.3 Results

3.3.1 Single environmental surface optimizations

The objective of the genetic algorithm is to identify the best parameter values to describe the relationships between individual landscape variables and gene flow. The transformation functions selected by the genetic algorithm for single surfaces largely support our hypotheses (**Table 3-1**). We identified a monotonically decreasing cost to movement with increasing pine volume and CSI values. Pine volume showed a decreasing relationship with cost that nears a linear relationship (**Figure 3-2-**B). In contrast, CSI showed a more marked plateau of high cost for low values of CSI. Thus, the rate at which CSI cost decreases is more pronounced at higher values of climate suitability (**Figure 3-2**-D). In other words, until a certain threshold in the suitability of climate for its survival and growth, cost to movement remains high for the MPB. Beyond this threshold, movement becomes easier and easier with increasing values of climate suitability. The negative effect of elevation is also captured with a transformation showing that the rate of elevation-cost increases is lower at higher elevations (**Figure 3-2**-A). Drought was fitted with a unimodal transformation with a maximum value at a drought value of 0, which indicates that the highest costs to beetle genetic connectivity are found in conditions with neither moisture deficit nor excess (**Figure 3-2**-C). The lowest costs are associated by severe drought (negative values).

3.3.2 Composite environmental surface optimizations – Integrated model

Building on these individual surfaces, we sought to identify an optimal composite cost surface that could be used to model MPB genetic connectivity. The results of the model optimization using log-likelihood selection (**Table 3-3**) indicate that a composite surface built from elevation and CSI costs creates the ecological distance with the strongest effect on genetic distance. Indeed, this surface had by far the largest Akaike weight and no other models had a comparable (difference ≤ 2) AICc (**Table 3-3**).

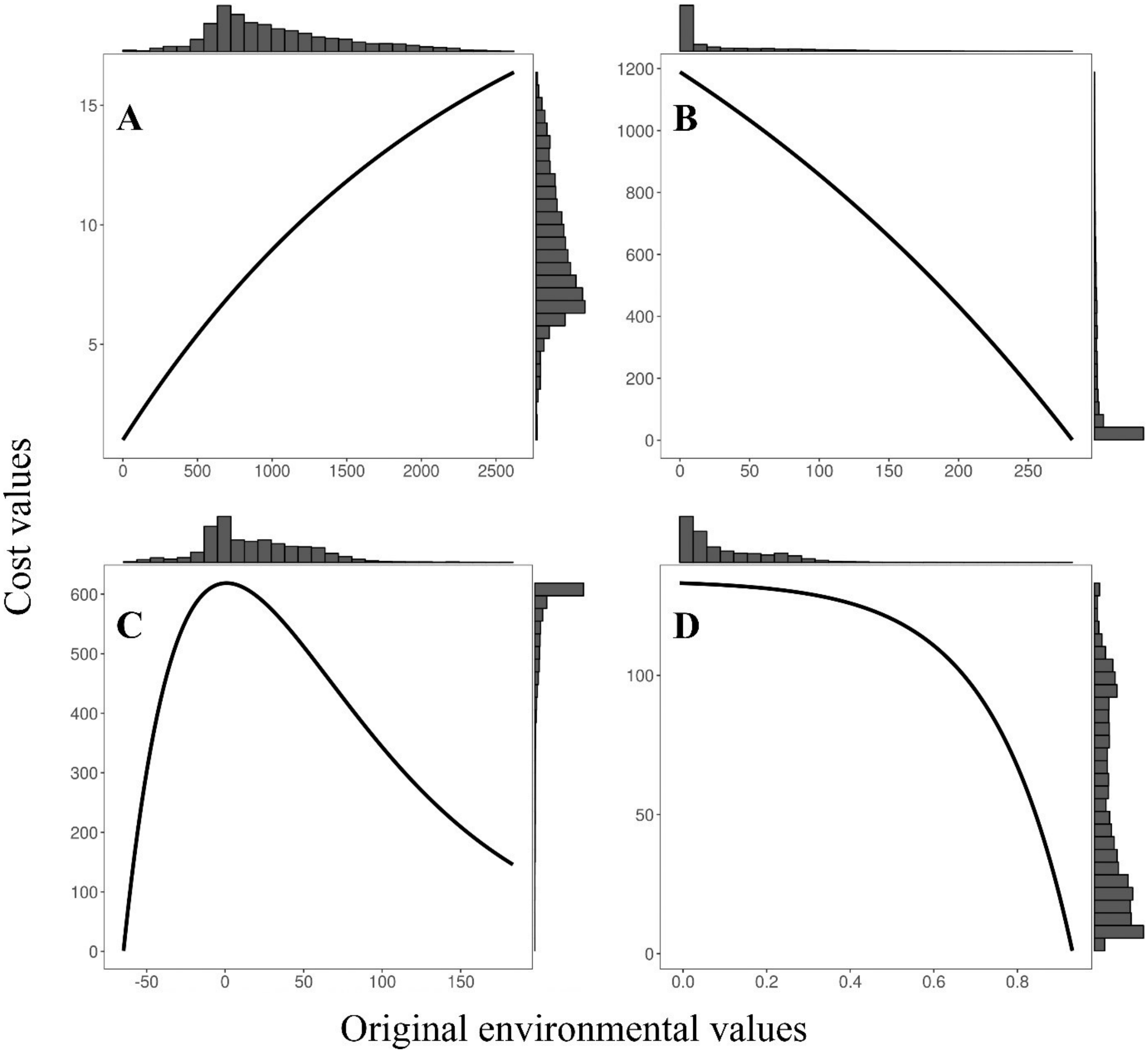


Figure 3-2 : Single-surface optimization results for each environmental variable: A) elevation, B) pine volume, C) drought, D) climate suitability index (CSI). The frequency distributions of cost values and original environmental values are shown on the sides of the graphs.

Results from the resampling analysis (**Table 3-3**) supported the likelihood-based inference and found that the elevation + CSI model best describes MPB genetic connectivity. Indeed, the elevation + CSI composite surface was ranked as the best model in 44.3% of resampling iterations, with an average rank of 2.72, making it the only surface that is, on average, among the top three models. The second most supported model was elevation + drought with 32.4% of iterations identified as the most supported model (**Table 3-3**). Rankings for other models diverge slightly between the full sample and resampling analyses. Akaike weights are more evenly distributed among models in the resampling analysis. Hence, support was attributed to more complex models with similar log-likelihoods to the best model in the full sample analysis in spite of penalties on complex models. However, out of those complex models, only elevation + drought + CSI achieved a high average rank and was regularly the top model in the resampling analysis (**Table 3-3**). The surface based on a homogeneous cost to movement (geographical distance) did not outperform any composite surface (**Table 3-3**). The best model for population connectivity in the northern cluster is the same as for the whole dataset: elevation + CSI.

3.3.3 Properties of the best surface: elevation + CSI

In the best surface, CSI and elevation costs contribute approximately 70% and 30%, respectively, of the total cost of travelling through a cell. The model using this surface as a predictor achieves a marginal R^2 of 0.67 and a conditional R^2 of 0.80, denoting a good performance. When visualizing the elevation + CSI composite surface that best explains genetic connectivity in our dataset (**Figure 3-3**), one can notice large areas with low costs in the north. The influence of the Rocky Mountains (high elevation) can be seen, with intermediate costs to movement covering much of the central part of our study area. The eastern part of our study area shows less heterogeneity in cost to movement, which is likely the result of reduced variation in elevation relative to BC and western AB (**Figure 3-3**).

Table 3-3 : Model selection and resampling analysis results sorted by Akaike weights. *Model* indicates the environmental cost surface(s) included in the optimized models. *LL* indicates the log-likelihood value of the model. *k* indicates the number of parameters used in each model. *AICc* indicates the AIC value of the model corrected for the number of parameters optimized (*k*) and the sample size. *ΔAICc* indicates the difference between the *AICc* of the model and the minimum *AICc* across all models. ω_{AICc} indicates Akaike weight of the model. *Rank* indicates the average rank achieved by the model. 1 is the top rank in each iteration. *Top %* indicates the frequency of the model reaching the top rank. Models with equal *Top %* values were then sorted by rank.

Model	LL	K	ΔAICc	ω_{AICc}	Rank	Top %
ELEVATION + CSI	919.13	5	0.00	0.94	2.72	44.3
ELEVATION + DROUGHT	915.67	5	6.90	0.03	3.48	32.4
ELEVATION + PINE + CSI	918.63	7	8.02	0.02	3.62	1.7
ELEVATION + DROUGHT + CSI	918.59	7	8.10	0.02	3.83	13.9
FULL MODEL	918.63	9	16.71	0.00	5.1	0.2
ELEVATION + PINE + DROUGHT	913.69	7	17.91	0.00	5.09	2.6
ELEVATION + PINE	908.55	5	21.16	0.00	7.45	0.5
PINE + DROUGHT	905.83	5	26.59	0.00	9.39	0.0
DROUGHT	903.25	3	28.71	0.00	10.12	0.0
PINE	902.63	3	29.94	0.00	12.24	0.1
DROUGHT + CSI	903.73	5	30.79	0.00	9.38	1.0
PINE + DROUGHT + CSI	905.53	7	34.15	0.00	10.5	1.2
PINE + CSI	902.03	5	34.19	0.00	13.1	1.2
GEOGRAPHICAL	897.66	2	34.58	0.00	13.74	0.1
ELEVATION	898.51	3	38.18	0.00	12.65	0.1
CSI	897.66	3	39.89	0.00	13.58	0.7
NULL	796.89	1	233.77	0.00	/	/

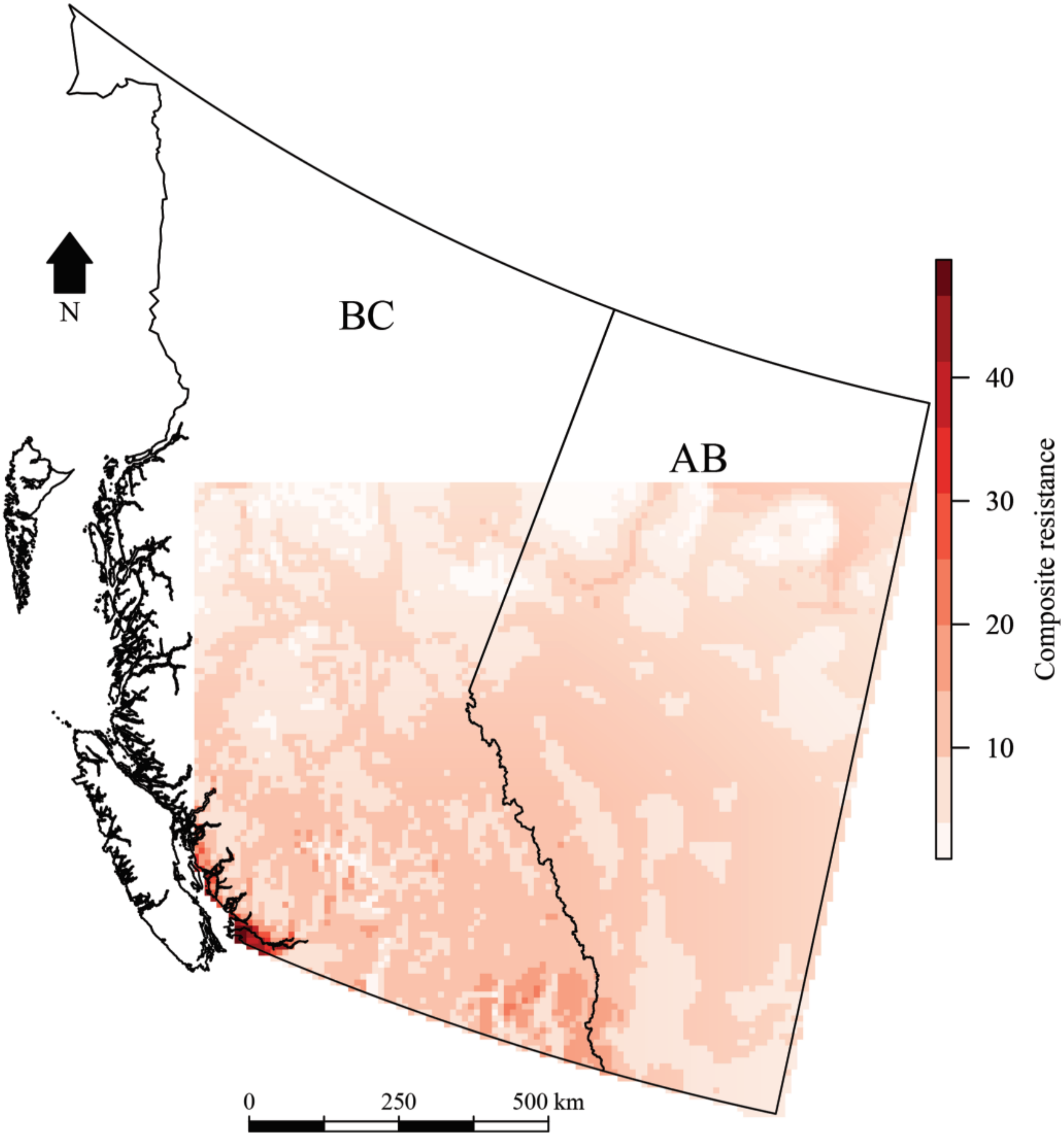


Figure 3-3 : Composite surface from our best model (elevation + CSI), projected on the study area. Lighter areas represent areas that offer less resistance and therefore facilitate movement.

3.3.4 Model application I - Origin of the Hinton MPB population

Based on our genetically informed, and machine-learning optimized model of landscape connectivity, we found that Jasper is the most likely source of beetles in Hinton, and that Hinton, Jasper and Edson are part of a connectivity corridor (**Figure 3-4**). Indeed, connectivity, as measured by commute-time distances was the highest between Hinton and Jasper (**Figure 3-4**, **Figure 3-4**). Connectivity was also strong between Edson and Hinton, with a difference of 7 451.9 relative to the Hinton-Jasper distance. Grande Prairie was rather isolated from Hinton according to our model, with a difference of 372 964.3 relative to the Hinton-Jasper distance.

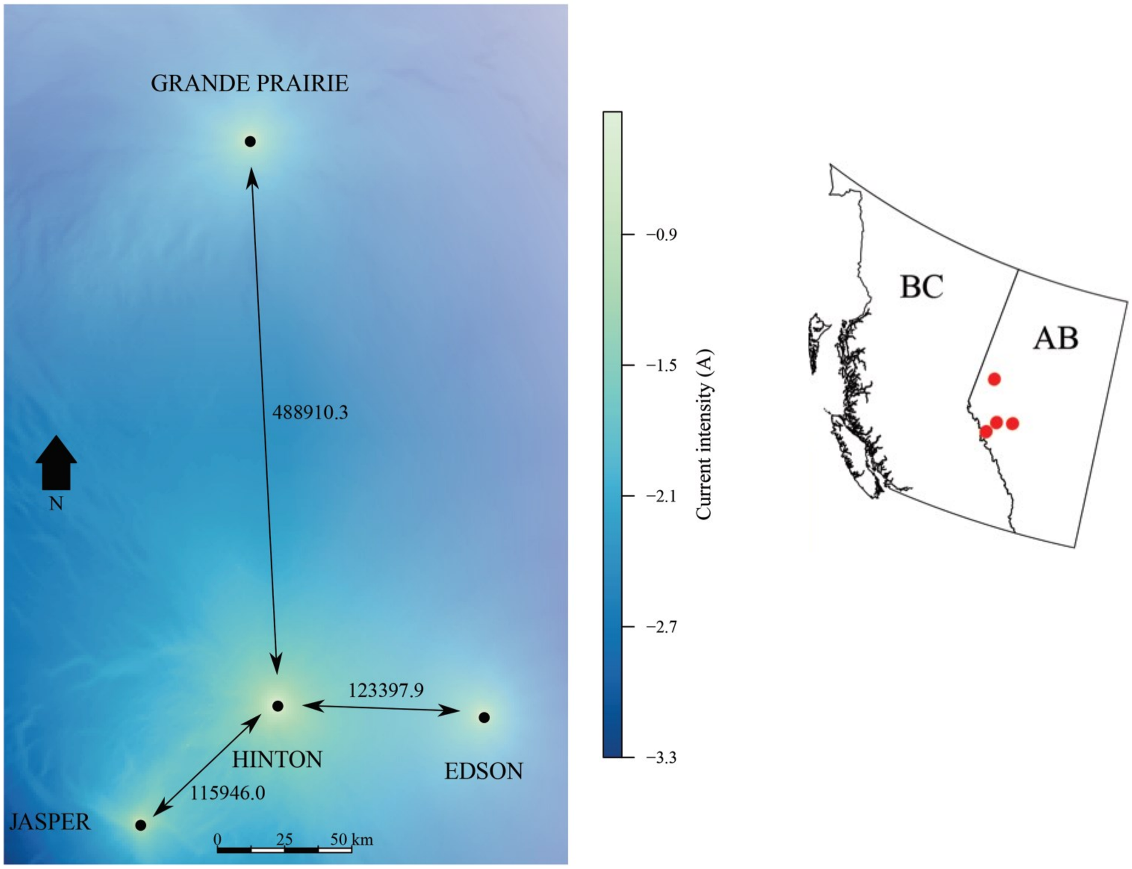


Figure 3-4 : Electrical current map based on the best model (elevation + CSI). The current intensity represents the expected concentration of movement between the populations: the lighter the colour or shading, the more movement we expect.

3.3.5 Model application II - Potential for eastward expansion

Predictions using both our final selected model and a pine-only model indicate that there are few obstacles to beetle movement between the 2008 outbreak limit and the eastern border of Saskatchewan (**Figure 3-5**). In predictions using our model based on elevation and CSI, there is higher connectivity in several large corridor areas: southeast of Wood Buffalo National Park (AB), southeast of Fort McMurray (AB), and a corridor crossing the Alberta and Saskatchewan border and following the southern limit part of the pine distribution (**Figure 3-5**-A). Prediction using a model based exclusively on pine volume (**Figure 3-5**-B) indicated that beetle connectivity is overall more homogeneous with fewer evident corridors, although one can recognize corridors identified using the elevation + CSI model.

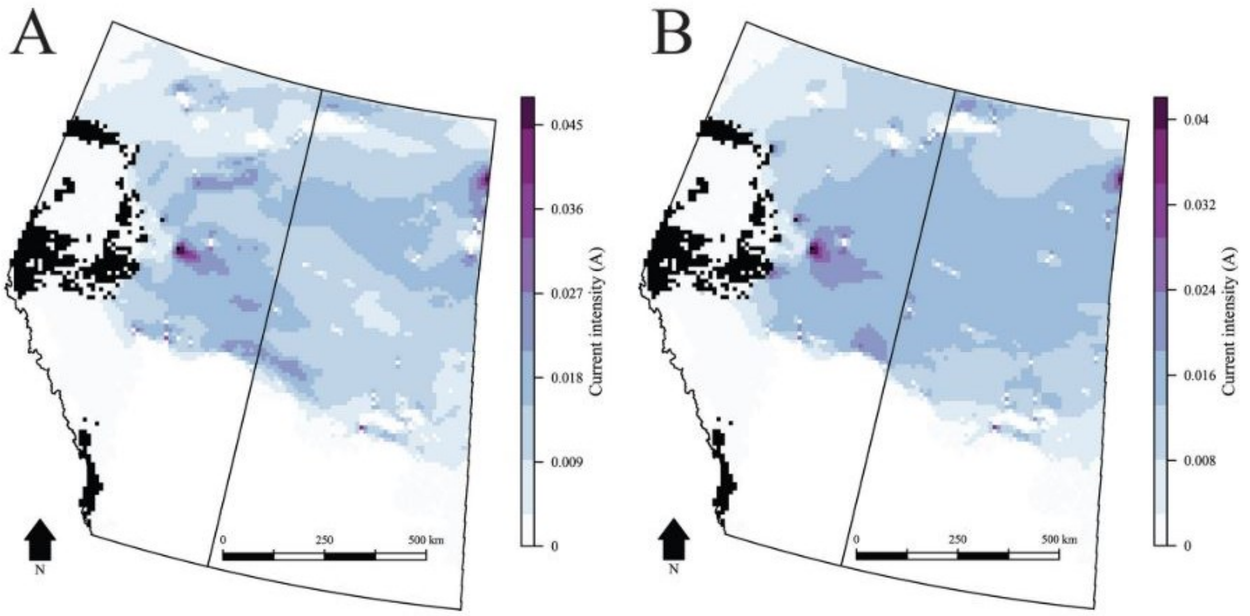


Figure 3-5 : Electrical current map based on an extrapolation to Alberta and Saskatchewan of A) our best model (elevation + CSI) and B) a model based on pine volume. The core infested area used as a source is represented in black. The electrical ground is the Saskatchewan–Manitoba border. The current intensity represents the expected concentration of movement between the populations: the darker the colour or shading, the more movement we expect.

3.4 Discussion

Improving our understanding of connectivity and movement is essential for the effective management of forest pest species. However, outbreaking populations pose many challenges such as expansion into novel habitats that limit applicability of expert opinion. We applied a novel machine-learning approach within a landscape genetics framework to characterize movement and population connectivity in the MPB, one of the most damaging forest insect pests in western Canada. Through single-surface models, we were able to support the hypotheses and describe how environmental heterogeneity translates into movement cost. Considering all combinations of these variables in addition to the single-surface models, we found pine beetle population connectivity is driven by a combination of elevation and climate. Using machine-learning tools that make no a priori assumptions about how landscape heterogeneity affects movement, we were also able to demonstrate the predictive strength of our MPB connectivity model in two management-relevant contexts.

3.4.1 Single environmental surface optimizations

The real value of our single-surface analyses goes beyond simply confirming the direction of landscape effects on gene flow (**Table 3-1**): it lies in the parameterization of the shape and the maximum value of those relationships. As hypothesized, elevation was negatively associated with gene flow: resistance increased with elevation (**Figure 3-2**-A). This negative association could be explained by higher mortality of pines due to beetle attack at lower elevation, enhancing reproduction of beetles and effective dispersal (Amman et al. 1973). Very high elevations are also associated with a decrease in climate suitability for the pine (Smithers 1961), which would affect the resources available to the MPB. Beyond an effect on reproduction, low-elevation valleys have been previously shown to facilitate beetle dispersal (de la Giroday et al. 2011). Although elevation emerged as an important predictor of beetle connectivity in our study area, given the lack of significant topography east of the Rocky Mountains, elevation will not likely be a significant factor influencing pine beetle outbreak spread.

Pine volume was positively associated with gene flow: resistance decreased with greater pine volume (**Figure 3-2**-B). High-volume stands are generally thought to be more susceptible to MPB attack (Safranyik et al. 2010a) and to result in higher beetle reproductive rates (Safranyik & Carroll 2006). Consequently, one expects that lower pine volume east of the Rocky Mountains, where the beetle is currently spreading, to constrain MPB spread.

The single-surface optimization of drought produced the only surface with a non-monotonic parameterization (**Figure 3-2**-C). The lowest costs to movement are associated with the driest and wettest conditions, whereas the highest costs are associated with intermediate levels of drought. High levels of drought are expected to facilitate movement and MPB population growth because drought-induced stress decreases the defences of mature host trees (Berg et al. 2006; Raffa et al. 2008). The fact that resistance to movement also decreases under the moistest conditions could be a statistical artefact associated with the complex edges of the coast that support the wettest areas (Koen et al. 2010) or be due to higher MPB brood production associated with increased water storage in pines, possible when excess water is available (Amman 1972; Safranyik & Carroll 2006). Taken independent of any other factors, drier conditions east of the Rocky Mountains might be expected to facilitate MPB spread.

Finally, we found that resistance decreased with higher values of climate suitability (**Figure 3-2**-D). Relative to the other monotonic relationships (e.g., elevation and pine volume), the effect of climate suitability on gene flow is less linear, with a plateau of high resistance followed by a sharper decrease of resistance than, for example, the pine single surface (**Figure 3-2**). Climate is known to strongly affect MPB (Bentz et al. 2010; Shore et al. 2004) at several stages of its life cycle. Indeed, favourable climate may enhance the synchrony of emergence and therefore facilitate overwhelming mass attacks on pines (Carroll et al. 2004; Logan & Amman 1991), increase the proportion of univoltinism relative to semivoltinism (Bentz & Powell 2014; Logan & Bentz 1999), and increase overall survival through easier overwintering (Bentz & Mullins 1999; Safranyik & Llnton 1998; Safranyik & Wilson 2006). The nonmonotonic relationship between climate suitability and resistance (**Figure 3-2**.-D) likely reflects the physiological limit of MPB at lower temperatures. The threshold after which the cost to movement based on CSI decreased sharply is very close to the value described by the creators of the CSI as the lower bound of the "extreme climatic suitability class" (Carroll et al. 2004), which denotes an agreement between ecological conductance and habitat quality for this variable. Replicated executions of the genetic algorithm, which ran without assumptions about the functional relationships between gene flow and landscape context, supported the signs of the relationships we hypothesized from the literature.

3.4.2 Composite environmental surface optimizations – Integrated model

Building a multisurface composite landscape genetics model and going beyond individually optimized resistance surfaces is important because using a single surface in a landscape genetics model assumes that only this variable influences genetic connectivity, which is rarely the case (Spear et al. 2010). A combination of elevation and CSI was the best model according to model selection which was corroborated by our resampling procedure. Elevation was consistently included in the best models, which indicates its importance to MPB gene flow.

Our results shed light on how the MPB outbreak may have exceeded its previous range. While topography has not changed during the last century, climate has (Masson-Delmotte et al. 2021). Previous research has shown that MPB outbreak dynamics in British Columbia have been driven mainly by increasing temperature during the last century (Raffa et al. 2008; Sambaraju et al. 2012; Shore et al. 2004). Our study builds on this previous research by accurately describing how climate affects the movement aspect of outbreak dynamics. Indeed, given the increasingly more suitable climate, our model may explain how the MPB was able to move so quickly east of the Rockies in recent times (Cooke & Carroll 2017; Janes et al. 2014).

Our results differ from a previous landscape genetics study on the MPB at the leading edge of the outbreak (James et al. 2011a). Using a neighbourhood-based analysis of landscape connectivity, in a central subset of our study area, James et al. (2011a) found that pine volume was an important predictor of connectivity for a northern genetic cluster of populations. Pine volume was not included in our best models. Our optimized models of how cost varies in response to both climate and elevation may have captured some of the effect of pine volume. Indeed, ecological distances calculated from the pine volume surface are strongly correlated to ecological distances calculated from the composite elevation and climate surface (Mantel correlation: 0.87; p-value $< 10^{-7}$). The costs to movement from both surfaces are weakly positively correlated (Pearson correlation: 0.18; p-value $< 10^{-16}$). This holds even when only considering the 75% most forested areas (Pearson correlation: 0.20; p-value < 10-16), to exclude unforested parts which cannot bear epidemic populations from the calculation. Therefore, if part of the variation in genetic connectivity explained by pine volume may be shared with climate and elevation, it is likely a minor part.

3.4.3 Model applications

Our first model application attempted to resolve a local-scale question about the likely origin of beetles found in Hinton, AB. Our connectivity analysis suggests that the beetles most likely originated from Jasper (**Figure 3-4**). Hinton is also well connected with Edson, and there seems to be a large corridor of high connectivity from Jasper to Edson. A recent study also showed that beetles east of Hinton are genetically similar to beetles from Jasper (Trevoy et al. 2018). Also, the latest MPB population forecast survey in Alberta showed that based on larval mortality, beetle numbers were strongly increasing in an area ranging from Jasper to the Edson forest area (*Mountain Pine Beetle Population Forecast Survey* 2017), which is similar to the highly conductive area that we described (**Figure 3-4**). From a management perspective, this model could be used to forecast connectivity between attacked and unattacked stands and to prioritize well-connected, but not yet attacked, stands for pre-emptive harvest.

Our second model application examined potential routes of eastward expansion at a larger, interprovincial scale (**Figure 3-5**-A). Through our exploratory analysis, there appears to be few constraints to beetle movement to the east when considering connectivity models, based on elevation and climate, or based on pine volume. Several large-scale high-connectivity corridors exist, notably through the southern part of the pine distribution (**Figure 3-5**), which has been previously described as conductive (Safranyik et al. 2010a). Although producing reliable models for predicting further expansion is challenging (Cooke & Carroll 2017), the low variation of the factors associated with MPB dynamics in the elevation + CSI model or the pine volume model sets the stage for future connectivity. According to the elevation and climate model or the pine volume model, managing the outbreak and mitigating its consequences would be difficult. Indeed, in the absence of localized, pinch-point areas of high connectivity, deciding in which areas to prioritize management efforts would be less straightforward and management efforts could be less efficient in slowing or reducing the consequences of an MPB outbreak.

3.4.4 Limits to our approach

An important consideration in the application of any spatial statistical model is that it can be difficult to reliably make predictions in areas outside the scope of the original data. The concern about model transferability, i.e., applying a model built using a spatial dataset to a subset of those data or to a different dataset, is widespread in ecological studies (Wenger & Olden 2012; Yates et al. 2018). A challenge with transferring models to new data is that predictions can be affected by many factors not necessarily included in the original model such as changing biotic interactions, sampling biases, and landscape dissimilarity between the original and the novel landscapes (Wenger & Olden 2012; Yates et al. 2018). This is especially true in studies dealing with species distribution (Moon et al. 2017; Petitpierre et al. 2017).

In the case of our model of MPB population connectivity, we do not know if the functional relationships between landscape heterogeneity and gene flow are the same outside our study area (**Figure 3-1**). For example, drought is likely to be more severe in the future in the continental zone of western Canada (Wang et al. 2014), which may ultimately reduce rather than increase the rate of MPB spread as moderate drought would (Amman 1972; Safranyik & Carroll 2006). This relationship, however, remains uncaptured by our models built using data from non-continental regions. Additionally, the cost associated with pine volume could change across the landscape as the main species of pine switches across western Canada, thereby changing biological interactions (Safranyik et al. 2010a). For example, MPB could move at a different pace in jack pine relative to lodgepole pine forests due to differences in reproduction and physiology. The uncertainty inherent with these new interactions justifies continued work on MPB connectivity, especially in novel habitats.

Validating wildlife landscape connectivity models is challenging because the processes behind connectivity are as diverse as they are complex to comprehensively observe. Often one must use indirect methods to infer connectivity as it is neither possible nor feasible to directly measure such processes. Were it possible, we would not need to use indirect methods such as genetic variation, which is complex to interpret back to actual individual movement. Indeed, relationships between gene flow and dispersal, can be complex and context dependent (Cayuela et al. 2018; van Strien 2017; Webster et al. 2002; Whitlock & McCauley 1999). Comprehensive validation is necessary because, although validation results are sometimes contradictory, they are used to support conservation actions (Correa Ayram et al. 2016; Laliberté & St-Laurent 2020; Wood et al. 2021). Fortunately, despite it being too rarely done, validation is often possible through various means. As in many modelling approaches, internal cross-validation can be used to evaluate how precisely predictions based on a model would perform on a yet-unseen dataset. While we tried to limit overfitting through a modified resampling procedure, we did not conduct a proper internal cross-validation. Internal validation of connectivity is sometimes done with a different model, such as logistic regression when the empirical data is discrete (Brennan et al. 2020). In our case however, such a proper internal cross-validation (e.g., leave one out) would have required the optimization of many more models and is generally considered unfeasible when using the method we used (Winiarski et al. 2020). In our case, running it would have required at least 81 CPU-years, which we did not have access to.

Empirical external validation of landscape connectivity models is also an option to verify the output of those models (e.g., current density), and justify their use to guide management efforts. For example, to validate a multispecies connectivity map created based on expert opinion about the resistance of various land cover features, Koen et al. 2014 used two empirical datasets. The first one is based on road mortality for 20 species of amphibians, snakes, and turtles, which should be positively related to movement routes for those species. In this case, it is assumed that more deaths are associated with lower resistance overall. The second one is based on telemetry data which was used to estimate home ranges. They assumed that areas with heavy space use should be associated with higher current density. Both datasets validated their connectivity model. A recent study which conducted similar modelling at an even larger scale and with more validation datasets provided further support (Pither et al. 2021). In another successful example, Laliberté & St-Laurent 2020 used an extensive dataset including four different empirical measures that were assumed to be associated with connectivity for two cervid species, to validate connectivity models built from a landscape featuring one major linear potential barrier.

Empirical validation of connectivity mapping techniques should be conducted in most studies, but it poses many new and complex challenges as well. While Laliberté & St-Laurent 2020 were able to successfully validate their landscape connectivity models for one species (likely because the telemetry data for the other species was too dated/inaccurate), they also highlighted that validation metrics largely differed in their association with connectivity model results. In essence, this may mean that studies rightfully aiming to validate their connectivity models with an external empirical proxy of movement, may need to “validate” their empirical validation metric before concluding about the usefulness of the connectivity models (Bond et al. 2017). This supplementary level of validation could be especially important when several validation metrics cannot be compared in the same study. Unfortunately, although it was recently spearheaded by Laliberté & St-Laurent 2020, studies using different validation metrics are yet a very rare occurrence in the already small subset of landscape connectivity modelling studies providing external validation. A potential solution could be to conduct spatially explicit simulations of gene flow (Landguth et al. 2017a) or individual movement/detection (Neilson et al. 2018b). Such simulations could then be used to estimate specifically how a validation metric is related to the original measure of connectivity provided by the modelling (i.e., least-cost paths, randomized shortest paths, circuit theory…). Analytical modelling approaches also vary in their relevance for specific applied objectives, and several should be compared, when possible (Bond et al. 2017; McClure et al. 2016; Poor et al. 2012; Zeller et al. 2020), otherwise the validation effort could be wasted.

Beyond choosing the right modelling approach and the right validation metric, there are more challenges concerning the external validation of landscape connectivity models. The relationship between validation metrics and connectivity may not be linear, or monotonic, and the distribution of the metric may be complex. This may force researchers to use ranks instead of the original values (Laliberté & St-Laurent 2020). This relationship may also depend on scale, species, season, and the algorithm used to model movement, making simulations even more valuable to understand this relationship. For example, the relationship between connectivity and wildlife road collisions may be ambiguous in some ecological contexts. Road mortality may be used as a proxy of the probability of movement across a road. However, for some species/populations, mortality may become such a cost to demography that its relationship with the permeability of the landscape changes to that of a barrier. In other words, potential permeability is often context-dependent and may not always lead to effective permeability. The positive relationship between habitat suitability/selection and connectivity that is assumed in the literature (Beninde et al. 2016; Koen et al. 2014; Laliberté & St-Laurent 2020; Walpole et al. 2012; Zeller et al. 2018), including this chapter, may not always hold. Indeed, conflating resource selection with connectivity sometimes leads to misleading parameterization of landscape resistance surfaces (Abrahms et al. 2017; Keeley et al. 2017). Furthermore, resource limitation and high densities may sometimes drive movement creating a negative relationship between habitat quality and functional connectivity (Larroque et al. 2022). Finally, one of the most counter-intuitive landscape connectivity modelling result (that was empirically validated) is that high probability of movement may not always be positively associated with gene flow (Marrotte et al. 2017). This can notably be explained by the fact that narrow corridors presenting very high current density (i.e., pinch points), hence high probability of movement, are surrounded by areas with high resistance making those pinch points overall dysfunctional regarding effective connectivity and gene flow (Marrotte et al. 2017). Hopefully, with the advent of open access animal movement databases (Kays et al. 2021) and genetic databases (e.g,, Dryad, GenBank, GGBN), validation will become a more common, accurate and reproducible step in landscape connectivity studies.

The huge landscape covered by MPB outbreaks, and their complex spatiotemporal dynamics, prevent replication and experimental manipulation, and makes using most external validation metrics difficult. The output of landscape connectivity models based on genetic data is often probabilistic, which contrast with most validation metrics used in the literature (e.g., for vertebrates), and optimizing model parameterization is computationally intensive, as described earlier. Despite those drawbacks, genetic data was our only choice to get an accurate picture of effective MPB dispersal. A recent study (Shegelski et al. 2021) provides support for one of our results. We predicted that Hinton would be colonized quickly, and that the colonizing beetles would originate from Jasper rather than from the northern, already infested, Grande Prairie area. Field surveys conducted by Alberta Environment and Parks showed that beetles were already present around Hinton in 2015 (ten years after the sampling used on our study). The vast majority of beetles collected around Hinton in 2016/2017 were most closely related to beetles sampled in the Jasper area (Shegelski et al. 2021). Beyond this coarse validation of our results about the geographic origin of the Hinton beetles, there is now also ample evidence that the spread was demographically intense, as expected in an outbreak. Indeed, the outbreak destroyed half of the total pine forest in Jasper in 2017, and the Hinton area was likely hit with a large in-flight of MPB in 2018 (Tom Daniels, West Fraser forestry superintendent, pers. comm., 2018). Monitoring flying insects is challenging (Osborne et al. 2002), but non-genetic data may also provide support for our landscape connectivity model, in our case, a posteriori. Aerial surveys of forest damage, field observations (including through pheromone-baited traps), and even weather surveillance radar (Ainslie & Jackson 2011; Bauer et al. 2019; Boulanger et al. 2017) may help us follow an outbreak. For example, a recent field survey generally supports our finding about the generally high connectivity at the southern edge of the forested area at the Alberta/Saskatchewan. Indeed, a recent statement by the Saskatchewan government has indeed reported that the MPB outbreak "has spread across Alberta to within 40 kilometres of the Saskatchewan border" (Government of Saskatchewan 2021).

While some studies may successfully represent biological complexity by making informed biological assumptions about the relationship between movement and landscape features, it is likely that expert opinion sometimes strays away from reality. This issue with biological realism may not always be avoided by carefully using traditional modelling practices. Indeed, a number of studies have stressed the limitations of expert opinion and incomplete direct observation in understanding movement, as well as issues with expert perception about non-linearity in resistance of the landscape to species movement.

Beyond statistical performance and parameter space considerations, one of the main advantages of the machine learning approach we used is to avoid the unfortunately common pitfalls of models based on expert opinion, which may sometimes very poorly conceptualize real biological relationships between a species and its environment. Empirical direct measurement of resistance to movement is quite rare and most experts would consider it a great resource to build landscape connectivity models. In one such study, experiments showed that forest habitat presented high resistance (Stevens et al. 2004) to juvenile natterjack toads (*Epidalea calamita*) movement. Based on this result, forested areas were expected to hinder toadlet dispersal. A behavioural study found that toadlets preferably choose forest over other habitats which facilitated movement relative to forest, when moving from the natal habitat (Stevens et al. 2006a). Landscape connectivity models parameterized using preference performed better than models based on slower movement through forests, which counter-intuitively led to the recommendation of preserving previously deemed unfavourable forest corridors (Stevens et al. 2006b). In another example highlighting the risk of making biological assumptions based on expert opinions and limited knowledge, field observations suggested that roe deer (*Capreolus capreolus*) move through valley bottoms, however, contrary to hypotheses based on expert opinion, an analysis based on telemetry data showed that roe deer avoided valley bottoms (Coulon et al. 2008). California tiger salamander (*Ambystoma californiense*) is a pond-breeding amphibian which is mostly fossorial and thus presents a cryptic life cycle. While the species is typically associated with grassland (Loredo et al. 1996; Trenham 2001) where it finds animal burrows to live a large part of its life in, a study unexpectedly found that the previously thought inhospitable chaparral habitat was only half as costly to move through as grassland (Wang et al. 2009). Not only did a coarse and simple optimization, and even a null model, perform much better than a model created by an expert panel in explaining the connectivity of several populations of spotted salamanders (*Ambystoma maculatum*) and wood frogs (*Lithobates sylvaticus*), but the parameter values produced by the panel and the optimization, and their ranks, were very different (Charney 2012). While capture-mark-recapture showed that short-distance movements of the threatened southern damselfly *(Coenagrion mercuriale*) were expectedly restricted to its natural habitat (streams), genetic estimates of dispersal showed that long-distance dispersal events happened over, much larger distances and through seemingly unfavourable habitat (Keller & Holderegger 2013; Keller et al. 2012). Focusing on the rate of water loss experienced by adult western slimy salamanders (*Plethodon albagula*), one study demonstrated an unexpected hump-shaped relationship with movement, with very favourable and very unfavourable habitats both presenting high resistance when a monotonic negative relationship was expected (Peterman et al. 2014). Roads and buildings unexpectedly presented low resistance to European pine marten (*Martes martes*) movement (Larroque et al. 2016), when it was previously widely described as avoiding human-dominated areas (Brainerd & Rolstad 2002; Storch et al. 1990).

From a biological point of view, methods based on optimization through machine learning may be particularly useful when expert opinion is not sufficiently documented about the non-linearity and non-monotony of the biological relationship between movement and landscape. Indeed, machine learning approaches to landscape connectivity modelling such as ResistanceGA (Peterman 2018) are able to pick up non-monotonic biological relationships (Larroque et al. 2022; Peterman et al. 2014; Wittische et al. 2019) which is often an admittedly difficult task based on expert opinion, notably for cryptic species such as fossorial species. Non-monotony is an especially important property because, although failing to accurately parameterizing absolute values may not be prohibitive, appropriately setting ranks (e.g., between levels in a categorical surface) is necessary (Bowman et al. 2020). The ResistanceGA approach was recently evaluated through extensive simulations and performed well in most scenarios (Winiarski et al. 2020) although there is room for improvement (Peterman & Pope 2021). Finally, another added difficulty which comes with non-linearity and the creation of composite surfaces, is that models based on expert opinion are often only exploring a limited portion of parameter space. This limitation means that artificial intelligence optimization approaches can explore a more diverse array of potential biological relationships between species movement and landscape resistance. Although it was out of scope for this chapter, comparing models based on expert opinion and those created through optimization and less supervision is nonetheless very interesting, and should be conducted regularly to better understand strengths, weaknesses and the potential complementarity of those approaches. A final important consideration is that we assume isotropic resistance to movement, i.e., the resistance between two locations is the same regardless of the direction being travelled; however, anisotropic movement and asymmetric resistance to movement and gene flow can play a significant role in shaping spatial patterns of genetic variation (Holderegger & Gugerli 2012). To our knowledge, no work has successfully incorporated directional processes explicitly in models of genetic connectivity, although recent efforts towards this goal have been made (Landguth et al. 2017c). Within the MPB system, wind (direction and speed) has been hypothesized influence the dynamics of long-distance dispersal (Ainslie & Jackson 2011; de la Giroday et al. 2012). We think that being able to use wind in combination with other isotropic landscape features would constitute the most well-rounded approach to modelling MPB movement. Incorporating directional processes such as wind in a landscape genetics model of the MPB is a natural, yet challenging, next step to better understanding the spatial dynamics of this complex system.

3.4.5 Conclusion

Using machine learning, we found that elevation and climate together constitute the best predictors of gene flow and movement of the MPB in western Canada. Using this model of landscape connectivity, we demonstrated that the beetles that recently colonized forests around Hinton, AB, most likely originated in Jasper and travelled in a high-connectivity corridor spanning from Jasper to Edson, in accordance with recent genetic analyses and government survival-based reports. This gives us confidence in the ability of our model to accurately predict population connectivity and gene flow within the spatial scope of the data that we used. Finally, using our results outside of the scope of the study area, we found that pine forests in eastern Alberta and Saskatchewan may provide homogeneous routes of colonization for the MPB if it continues its expansion. Future avenues of research include incorporating directional processes such as wind into the landscape genetic model and evaluating uncertainty in long-term forecasts of the spread of MPB populations in novel habitats.

Acknowledgements

This research was supported by a grant to PMAJ and the TRIA Network from the Natural Science and Engineering Research Council of Canada (grant no. NET GP 434810-12), with contributions from Alberta Agriculture and Forestry, fRI Research, Manitoba Conservation and Water Stewardship, Natural Resources Canada - Canadian Forest Service, Northwest Territories Environment and Natural Resources, Ontario Ministry of Natural Resources and Forestry, Saskatchewan Ministry of Environment, West Fraser and Weyerhaeuser. JW was also supported by a scholarship from the Forest Complexity Modelling (FCM) NSERC CREATE. We thank William Peterman for his advice regarding the use of *ResistanceGA* and Jeremy Larroque for help with GIS work. We thank David Coltman and Felix Sperling for their suggestions about our first model application. Finally, we thank Érik L'Heureux, Simon Legault, and Felix Sperling for their comments on an earlier version of the manuscript.

Supplementary material

Supp. Table 3-1 : F_{ST} values for all pairs of populations listed in **Table 3-2**.

POP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
1	0																										
2	0.013	0																									
3	0.018	0.013	0																								
4	0.022	0.027	0.039	0																							
5	0.041	0.05	0.065	0.022	0																						
6	0.054	0.062	0.079	0.038	0.02	0																					
7	0.039	0.047	0.061	0.021	0.03	0.045	0																				
8	0.027	0.021	0.017	0.055	0.084	0.101	0.077	0																			
9	0.025	0.019	0.014	0.051	0.078	0.095	0.073	0.017	0																		
10	0.025	0.02	0.013	0.051	0.081	0.097	0.076	0.016	0.014	0																	
11	0.025	0.021	0.015	0.051	0.079	0.096	0.076	0.018	0.014	0.014	0																
12	0.02	0.017	0.012	0.046	0.073	0.089	0.069	0.014	0.014	0.013	0.014	0															
13	0.022	0.017	0.012	0.047	0.075	0.09	0.07	0.013	0.013	0.013	0.013	0.011	0														
14	0.023	0.018	0.014	0.049	0.076	0.092	0.071	0.015	0.014	0.014	0.015	0.012	0.011	0													
15	0.022	0.018	0.014	0.048	0.075	0.091	0.071	0.013	0.013	0.016	0.014	0.012	0.011	0.012	0												
16	0.023	0.018	0.013	0.049	0.076	0.093	0.072	0.013	0.014	0.014	0.015	0.012	0.012	0.013	0.012	0											
17	0.029	0.037	0.05	0.015	0.022	0.037	0.02	0.067	0.062	0.064	0.062	0.057	0.059	0.06	0.061	0.061	0										
18	0.042	0.05	0.062	0.026	0.033	0.049	0.021	0.077	0.073	0.076	0.083	0.068	0.066	0.074	0.069	0.072	0.025	0									
19	0.046	0.056	0.069	0.025	0.035	0.051	0.015	0.086	0.083	0.085	0.084	0.078	0.078	0.08	0.08	0.081	0.023	0.018	0								
20	0.046	0.055	0.069	0.025	0.035	0.051	0.015	0.085	0.081	0.084	0.083	0.076	0.077	0.079	0.077	0.08	0.024	0.018	0.014	0							
21	0.039	0.047	0.059	0.02	0.031	0.045	0.014	0.076	0.072	0.074	0.073	0.067	0.068	0.069	0.069	0.07	0.018	0.02	0.014	0.015	0						
22	0.014	0.013	0.014	0.029	0.052	0.066	0.05	0.02	0.017	0.018	0.019	0.016	0.016	0.018	0.017	0.017	0.04	0.051	0.057	0.057	0.048	0					
23	0.041	0.05	0.062	0.021	0.031	0.045	0.016	0.079	0.076	0.078	0.078	0.07	0.072	0.074	0.073	0.074	0.02	0.018	0.014	0.015	0.015	0.051	0				
24	0.054	0.063	0.077	0.03	0.042	0.056	0.021	0.094	0.091	0.092	0.093	0.085	0.086	0.089	0.087	0.088	0.029	0.026	0.019	0.02	0.02	0.064	0.02	0			
25	0.015	0.013	0.015	0.031	0.054	0.067	0.051	0.023	0.021	0.019	0.02	0.016	0.017	0.02	0.018	0.018	0.039	0.053	0.059	0.057	0.048	0.015	0.052	0.065	0		
26	0.017	0.015	0.013	0.039	0.062	0.077	0.06	0.019	0.016	0.017	0.019	0.016	0.015	0.015	0.015	0.016	0.05	0.065	0.069	0.068	0.059	0.015	0.062	0.076	0.016	0	
27	0.014	0.019	0.027	0.014	0.028	0.042	0.025	0.038	0.036	0.037	0.037	0.032	0.032	0.034	0.034	0.034	0.018	0.029	0.031	0.031	0.024	0.02	0.027	0.038	0.021	0.026	0

Supp. Table 3-2 : Commute-time distances for all pairs of populations listed in **Table 3-2**.

POP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
1	0																										
2	43.69	0																									
3	37.1	39.95	0																								
4	40.09	51.55	46.73	0																							
5	55.44	66.72	57.84	59.27	0																						
6	48.24	58.83	46.3	55.88	54.17	0																					
7	63.04	71.65	66.05	62.89	69.07	72.41	0																				
8	69.51	68.32	65.28	74.08	89.17	82.89	90.45	0																			
9	58.02	56.48	47.47	65.42	78.97	70.18	83.96	72.83	0																		
10	64.01	62.23	54.62	71	84.85	76.45	89.32	75.85	44.6	0																	
11	79.38	77.69	69.59	86.51	100.3	91.72	104.9	92.18	57.3	50.66	0																
12	50.42	50.21	47.63	54.02	69.3	63.84	69.09	46.57	57.65	61.4	77.5	0															
13	59.6	61.49	58.86	61.52	77.04	72.92	74.41	70.49	71.64	76.04	91.93	45.63	0														
14	54.2	51.86	46.03	60.88	75.05	67.02	79.1	64.88	47.11	50.23	66.57	50.3	65.17	0													
15	59.34	59.42	57.25	62.54	78.18	72.94	77.83	65.78	68.49	72.58	88.57	42.24	52.35	61.5	0												
16	47.81	46	43.13	52.79	67.87	61.25	69.84	47.86	50.41	53.5	69.81	33.66	51.6	42.07	46.33	0											
17	48.76	59.54	52.46	49.94	48.03	55.22	53.93	80.79	72.09	77.75	93.22	60.39	67.26	67.75	69.15	59.72	0										
18	67.04	73.19	68.94	66.78	79.17	78.67	66.08	88.86	84.98	90.01	105.7	66.24	69.58	79.55	75.09	68.85	66.77	0									
19	82.16	88.41	84.03	82.14	93.49	93.47	77.63	104	100.1	105.2	120.9	81.29	84.8	94.72	90.31	84.03	80.8	69.11	0								
20	71.42	78.54	73.73	71.29	80.89	82.06	60.17	95.36	90.5	95.65	111.3	73.18	77.35	85.29	82.06	75.14	67.4	62.42	60.81	0							
21	64.99	71.67	67.28	64.24	76.9	76.58	64.69	88.33	83.72	88.83	104.5	66.17	69.99	78.41	74.81	68.09	64.33	54.51	76.4	66.59	0						
22	51.27	50.97	50.02	54.39	70.96	65.51	71.75	65.59	62.87	67.42	83.26	45.11	54.07	56.45	51.21	44.2	62.16	70.47	85.88	76.97	69.5	0					
23	66.15	72.69	68.35	65.48	78.21	77.79	66.11	89.14	84.69	89.78	105.4	66.89	70.59	79.36	75.55	68.95	65.72	53.46	76.6	67.26	31.6	70.37	0				
24	93.17	98.52	94.55	93.4	106.1	105.1	94.28	112.7	110	114.8	130.6	89.33	92.25	104.3	98.62	93.03	94.19	81.84	83.53	86.25	89.15	95.04	89.14	0			
25	49.82	49.56	49.58	52.94	70.4	64.79	72.13	69.36	63.89	68.8	84.52	49.52	58.63	58.01	57.09	47.69	61.8	71.96	87.39	78.04	70.59	42.31	71.54	97.04	0		
26	50.52	45.63	44.01	57.23	71.87	64.07	75.93	66.19	54.4	59.41	75.12	49.39	62.77	47.29	59.46	43.08	64.49	76.56	91.76	82.24	75.33	52.79	76.28	101.5	53.54	0	
27	49.67	53.51	51.56	49.55	68.86	64.52	69.26	72.73	67.17	72.23	87.91	52.08	58.7	61.67	59.63	51.61	59.47	69.07	84.76	75.26	67.27	48.79	68.28	94.71	45.96	57.85	0

Supp. Table 3-3 : Geographic distances for all pairs of populations listed in **Table 3-2**.

POP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
1	0																										
2	151.5	0																									
3	194.8	186.3	0																								
4	109.1	209.7	302.8	0																							
5	281.9	432.1	386.6	290.4	0																						
6	196.8	321.1	225.5	268.6	170.8	0																					
7	361.8	491.2	543.1	282.3	264.5	390.7	0																				
8	548.4	405.5	549.9	558.1	828.9	726.2	816.6	0																			
9	420.3	306.7	268.7	508.8	654.3	493.8	781.8	429.9	0																		
10	518.6	390.2	379.4	598.7	763.3	604.9	877.3	399.9	112.3	0																	
11	570.7	450.9	415.5	656.7	802	638.5	931.9	468.1	150.7	73.7	0																
12	555.5	433.2	604.2	537.2	823.9	747.5	766.7	141.3	538.5	526.6	597.6	0															
13	448.2	401.4	584.7	376.7	657.4	639.4	527	399.7	648.4	684.5	757.7	289.5	0														
14	372.3	226	295.2	433.8	643.3	505.5	716.1	282.2	160.7	189.1	260.8	379.9	498.5	0													
15	410.4	320.9	506.6	370.6	660.9	607.2	584.7	268.7	526.9	553.2	626.8	183.1	138.1	371.1	0												
16	421	276.4	424.3	438.6	702.3	597	706.7	129.2	342.4	344	417.4	197.8	357.9	182.5	220.2	0											
17	235.4	382	397.9	195.1	133.5	233.7	157.1	753	650.3	752.1	801	729.6	537.2	606.6	557.1	632	0										
18	443.6	528	638.5	336.7	452.4	543.4	203.3	760.1	834.7	913.8	977.7	678.8	401.8	733.1	498.9	672.3	324.3	0									
19	547.3	653.4	738.6	449	475.8	603.8	214.3	911	957.3	1043	1104	831.1	552.5	868.1	651.1	819.7	370.2	152.3	0								
20	493.5	611.9	680	402.9	395.6	531	140.8	901	910.5	1002	1059	832.7	565.7	832.6	650	801.2	297.7	170.8	84	0							
21	368.8	456.8	563.6	262.2	393.4	472.9	167.8	711.2	763.1	844.6	907.4	640.2	376.9	666.2	457.6	616.3	261.7	74.8	204.2	192.5	0						
22	265.4	195.5	379.9	229.6	518.5	462.1	471.2	345.8	459.6	513.6	584.1	308.1	205.9	324.8	145.6	244.3	421.6	429.1	575.5	558	372	0					
23	373.8	459.5	568.6	266.7	401.6	480.1	175.6	709.3	766	846.7	909.9	636.9	371.8	667.7	454.5	615.5	269.7	70	204.4	196	8.3	371.2	0				
24	848.6	930.3	1043	742.5	802.3	927.4	540.6	1101	1237	1310	1377	991.1	703.5	1125	832.8	1037	694.8	405.8	326.7	407.3	480.2	805.7	475.8	0			
25	201	145.7	324	173.8	458.9	397.8	433.8	386.8	434.2	500.4	567.9	365.1	261.4	315.4	209.6	272.9	368.6	419.1	559	532.6	355.1	64.4	355.7	810.1	0		
26	253.7	102.2	239.1	303.8	533.5	414	585.9	315.5	250.2	312.8	380.2	366.2	403.9	132.1	294.7	187.3	483.1	602.4	736.1	700.7	534.7	209.5	536.3	997.4	187.8	0	
27	195	194.1	353.9	130.2	420.4	385.4	367.6	449.4	495.4	567	633.1	412.7	254.2	384.1	241.2	340.5	317.3	349.2	488.1	462.4	284.3	105.9	285.1	743.5	70.8	254.3	0

4. CHAPITRE 4 – Détection des changements temporels de diversité génétique : un nouvel outil pour les études d'écologie moléculaire avec des enquêtes répétées

Detecting temporal changes in genetic diversity: a new tool for molecular ecology studies with repeated surveys

Julian Wittische[1], Pierre Legendre[1], Patrick M. A. James[1,2]

[1] Département de Sciences Biologiques, Campus MIL, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal, QC, Canada, H3C 3J7

[2] Institute of Forestry and Conservation, University of Toronto, 33 Willcocks St., Toronto, ON, Canada, M5S 2J5

Submission: This manuscript has been submitted to *Molecular Ecology Resources* on 4 October 2021.

Abstract

Understanding spatiotemporal changes in biodiversity, including genetic diversity, is essential to track the effects of global change and to inform effective conservation plans. Although temporal questions are common in community ecology, they are less often investigated in landscape genetics. Indeed, detecting changes in local genetic diversity due to demographic events that exceed what one would expect under neutral conditions (e.g., drift) is challenging. Our capacity to detect such changes is also information dependent. Existing methods to detect meaningful genetic changes through time typically require large datasets containing information beyond simple allele counts, such as mutational models and fine-scale population structure. However, when such extensive information is not available, methods are still needed to detect and understand temporal changes in genetic diversity induced by demographic events. In this paper, we describe Temporal Genetic Indices (TGI), a new method for identifying significant changes in genetic diversity through time. This method uses permutations of genotypic matrices to test the significance of genetic temporal change at sites, given neutral genetic change at other sampling sites in the study landscape. TGI overcomes existing challenges to detecting temporal change in genetic data with minimal information. We demonstrate the utility of TGI for identifying the genetic legacies of important historical demographic events using demo-genetic simulations. We further demonstrate the ability of our TGI approach to identify such legacies under different levels of dispersal, spatial extent of the demographic events, and the timing of sampling relative to the events. Finally, we successfully apply TGI to an empirical dataset, with our application providing a straightforward test for genetic change and supporting previous conclusions about the data. An R function to implement the method is now available, as well as utility functions for those wishing to further simulate and analyze their simulations.

Keywords: Time; Significance testing; Population genetics; Monitoring; R software

4.1 Introduction

Global biodiversity at the gene, species, population, and ecosystem scales is being altered at an increasing rate, with significant consequences for ecosystem functioning and the long-term viability of the biosphere (Bellard et al. 2012; Dirzo et al. 2014; Leigh et al. 2019). Given these global changes and the inherent temporal variability of biological systems, conservation biologists are increasingly recognizing that it is no longer sufficient to study spatial patterns in biodiversity at a single point in time. Instead, trends in biodiversity must be observed across both space and time (Allendorf et al. 2010; Bradburd & Ralph 2019; Fenderson et al. 2019). Important events in a population's history can be detected using genetic data (Bradburd & Ralph 2019), however, novel techniques are needed to quantify and track such events through both space and time.

Spatiotemporal variation in genetic diversity can provide important insights into the connectivity and demographic history of populations (Draheim et al. 2018; Moraes et al. 2017). Indeed, population genetics has proven essential for translating observed genetic variation into meaningful inferences that can inform conservation efforts (Allendorf et al. 2010; Harrisson et al. 2014; Segelbacher et al. 2010), and the causes and consequences of temporal variation in genetic diversity are at the crux of many conservation and public health issues (Díez-del-Molino et al. 2018; Lauterjung et al. 2019; Moraes et al. 2017). Researchers commonly explore patterns in spatiotemporal population genetic data (Banks et al. 2013) to quantify isolation-by-distance (Rousset 1997; Wright 1943), time since population bottlenecks (Gattepaille et al. 2013; Maruyama & Fuerstt 1985), rates of migration between isolated populations (Bezemer et al. 2019; Buschbom et al. 2011), and the timing and extent of outbreak expansions (James et al. 2015; Larroque et al. 2019; Wittische et al. 2019).

However, new approaches are needed to detect meaningful temporal variation in genetic diversity. The significance of genetic changes may be derived from the comparison to an expected reference distribution, or a null model based on typical background processes (e.g., drift, local gene, flow, mutation). Detecting such meaningful temporal changes represents a first step in elucidating the processes that govern demographically dynamic systems such as those found during population outbreaks (Fisher & Garner 2020; Maynard et al. 2017), major weather events (Poff et al. 2018), species invasions (Mack et al. 2000), or other disturbances such as a wildfire (Suárez et al. 2012). Temporal genetic analyses could similarly identify which populations, among a set of previously sampled populations, received migrants from long-distance dispersal events (Apodaca et al. 2013). Because temporal genetic variation reflects the evolutionary potential of a population and the probability of its persistence (Aeschbacher et al. 2017; Bolnick & Nosil 2007; Kremer et al. 2012), relating temporal genetic variation to landscape change can provide important insights about the eco-evolutionary dynamics of a species, and be used to inform conservation strategies (e.g., Landguth et al. 2017b).

There are currently two general approaches for investigating temporal genetic variation. The first suite of approaches uses statistical models to infer demographic history from genetic data obtained at a single time point (Excoffier et al. 2013; Gutenkunst et al. 2009; Kamm et al. 2020). This approach is often computationally intensive, requires high-quality microsatellite or extensive SNP datasets. This approach also requires extensive knowledge of the biological system, including information on recombination processes (Gattepaille et al. 2013) and ascertainment bias (Albrechtsen et al. 2010; Clark et al. 2005; Marth et al. 2004). The second suite of approaches compares genetic diversity between samples taken from the same sites over time using any genetic markers through either qualitative comparison or statistical models, ideally using a null reference distribution. These repeated-survey approaches are more readily usable in systems where less information is available, such as non-model species. Repeated-survey approaches can also be used in systems that were sampled in the past, with a goal of comparing contemporary to historical patterns (Moraes et al. 2017).

Despite our ability to compare genetic diversity at two points in time, several technical and conceptual challenges remain. One such challenge is determining which metric to use to meaningfully quantify and detect temporal changes. Some studies have used genetic differentiation metrics such as Jost's D or F_{ST} or its analogues (Knight et al. 2018; Larroque et al. 2019; Segura-García et al. 2019) to evaluate temporal changes between genetic datasets. However, translating our spatial understanding of these genetic differentiation indices to the temporal dimension is not straightforward (Bhatia et al. 2013). An additional challenge for temporal genetic analyses is disentangling spatial from temporal effects, because the additivity of genetic drift means that genetic differentiation can be associated with both space and time (Murray et al. 2016; Skoglund et al. 2014). Finally, repeated-survey analyses remain challenging because we lack sufficiently developed tools to distinguish natural temporal variation in genetic structure due to demographically induced drift, local gene flow, and mutation, from the changes caused by external forces. In other words, we need to develop an adequate distribution of reference patterns of genetic variation based on common background processes prior to deciding whether a change is meaningful.

Although its objectives differ from those of population genetics, the field of community ecology has a history of explicitly examining change in community composition through time. Temporal beta-diversity indices (TBI; Legendre 2019) are used to quantify and assess temporal changes in ecological community composition using a dissimilarity index calculated between samples taken at different times and at several sites. The significance of these dissimilarities is then tested using a permutational procedure. The temporal beta-diversity indices approach has effectively demonstrated temporal variation in simulated community composition (Legendre 2019) and in dozens of empirical datasets, but the potential of a temporal beta-diversity indices-inspired tool to detect meaningful temporal changes in genetic diversity has not yet been examined. Given the conceptual similarity in data structure between species diversity in multi-species community composition data, and genetic diversity in multilocus genotype data, we sought to determine how temporal beta-diversity indices could be modified to identify significant variation in spatiotemporal genotypic data.

In this paper, we propose and evaluate a method for extending the temporal beta-diversity indices framework to spatiotemporal population genetic data. Our new framework, which we call temporal genetic diversity indices (TGI), is designed to identify significant temporal variation in spatial genetic diversity using relatively information-poor genetic data while accounting for confounding forces such as drift, local gene flow, and mutation by creating a reference distribution. We demonstrate the effectiveness and applicability of the TGI approach using simulated genetic data, where each simulation combined multiple scenarios in which portions of a landscape were affected by a non-selective demographic change.

Specifically, we assess TGI's capacity to detect significant temporal variation in genetic diversity in three demographic factors: 1) species dispersal ability; 2) the number of populations affected by a demographic event (i.e., spatial extent of the event); and 3) the time between consecutive surveys. We predict that TGI's ability to detect temporal genetic changes will be lower in populations with higher dispersal capacity because of the homogenizing effect of higher gene flow. We also predict that our ability to detect changes will decrease as the time between successive sampling events increases. Beyond the method to simulate the null distribution, the selection of the cut-off value may also influence the performance of our permutational significance testing. Indeed, selecting a proper significance threshold for the *p*-value calculated from the TGI test permutations, and therefore defining which changes in genetic diversity are significant or not, is important for balancing selectivity and power. Therefore, we describe how this potential trade-off occurs along a range of thresholds and discuss its implications. Finally, we illustrate how TGI provides a functional testing framework by applying it to a real genetic dataset representing a large landscape with many populations of the Northern tidewater goby, a threatened species of fish found in western North America. A better understanding of the factors that influence changes in genetic diversity through time and improved techniques to monitor these changes is essential to describe and understand global biodiversity losses in the context of the current sixth mass extinction. That is why we built a tool to make meaningful assessments of the significance of temporal genetic diversity changes relative to an informed null model based on common background processes.

4.2 **Methods**

4.2.1 Adapting temporal beta-diversity indices for genetic data

Temporal beta-diversity indices (TBI) are calculated by computing dissimilarities in species composition between data surveyed at two different times for all sampling sites. Temporal beta-diversity indices give local measures of the change in community composition at each site; significance of these indices is then tested through simultaneous permutations of the two site-by-species input matrices. To extend temporal beta-diversity indices to TGI, we substituted community dissimilarities with genetic distances calculated from site-level allele frequencies in order to compare two different temporal surveys (see **Table 4-1**-A) for a simple example showing how we transformed two temporal surveys into a genetic distance, for a two-site landscape). The null hypothesis in this case is that genetic composition between both time points does not differ more than would be expected due to background processes typical to the landscape.

Neutral spatial processes such as genetic drift can produce temporal variation in spatial genetic structure. In designing TGI, our goal was to determine how to identify temporal changes that are significantly different from what would be expected under a scenario with drift, local gene flow, and mutation. Because there are no reference criteria for the changes in genetic diversity that would constitute a significant temporal genetic change, we used a permutation-based approach to generate a distribution of genetic distances to which an observed genetic distance can be compared.

Table 4-1: Examples of A) the computation of the original TGI values for a biallelic marker (the method works for microsatellites too!) and B) the way we permutated input genotypic matrices to create a distribution to test TGI significance. The numbers in the table represent the number of copies of an allele in a sampled population.

A)

	Locus	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
	Allele	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
FIRST SURVEY T_1	**Pop. 1**	4	0	4	0	4	0	3	1
	Pop. 2	0	4	1	3	2	2	3	1
	Pop. 3	1	3	1	3	0	4	0	4
	Pop. 4	0	4	0	4	1	3	1	3
SECOND SURVEY T_2	**Pop. 1**	2	2	3	1	3	1	2	2
	Pop. 2	3	1	1	3	1	3	2	2
	Pop. 3	1	3	2	2	1	3	0	4
	Pop. 4	0	4	0	4	2	2	1	3

TGI value for Population 1 between T_1 and T_2 (see Supp. Mat. 3-1):

$$TGI(pop_1) = D_{Roger}(pop_{1,\mathrm{T}_1}, pop_{1,\mathrm{T}_2}) = \frac{1}{4} \times \left(\left(\frac{4}{4} - \frac{2}{4} \right) + \left(\frac{4}{4} - \frac{3}{4} \right) + \left(\frac{4}{4} - \frac{3}{4} \right) + \left(\frac{3}{4} - \frac{2}{4} \right) \right)$$

$$TGI(pop_1) = \frac{1}{4} \times (0.5 + 0.25 + 0.25 + 0.25) = 0.3125$$

B)

Locus	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
Allele	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
Pop. 1	4	0	4	0	4	0	3	1
Pop. 2	0	4	1	3	2	2	3	1
Pop. 3	1	3	1	3	0	4	0	4
Pop. 4	0	4	0	4	1	3	1	3

Permutation example: this is done for both surveys and repeated 999 times to create the test for TGI. Permutation index 2, 4, 1, 3.

Locus	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
Allele	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
Pop. 1	0	4	1	3	2	2	3	1
Pop. 2	0	4	0	4	1	3	1	3
Pop. 3	4	0	4	0	4	0	3	1
Pop. 4	1	3	1	3	0	4	0	4

Table 4-2 : Two-factor simulation experiment with scenario abbreviations used throughout the manuscript. The third column indicates the number of affected populations with spatiotemporal population genetic legacies; scenarios with zero affected populations represent the control scenarios. We ran 180 unique simulations (replicates) for each combination of factor levels, which amounts to 2160 simulations in total.

Number	Dispersal	Affected populations	Code
1	Low	1	L1
2	Moderate	1	M1
3	High	1	H1
4	Low	2	L2
5	Moderate	2	M2
6	High	2	H2
7	Low	3	L3
8	Moderate	3	M3
9	High	3	H3
10	Low	0	CL
11	Moderate	0	CM
12	High	0	CH

We thus created a large number of randomly generated distribution of genetic distances. This allowed us to approximate a large number of landscapes which were generally genetically similar to our originally simulated landscape with the main difference being the diluted effect of the demographic event. While those randomly generated landscapes are created through permutations of genotypic matrices which erase the signal of the demographic event, not through simulations, their genetic variation can still be considered to retain the effects of many generations of drift, local gene flow, and mutation, created during the simulations (see section 4.3.2 for details about the simulations). For each of the two input genotypic matrices representing the two temporally distinct surveys containing data about the same alleles, loci and sites, we permuted the allele counts at each locus (**Table 4-1**-B). Permutations were performed using the *poppr* R package (see *Software*) to maintain allelic structure and heterozygosity (Agapow & Burt 2001). We used 999 random permutations in all analyses. To be clear, we created an adequate null model for testing purposes by combining realistic simulations of drift, local gene flow, and mutation with a permutation which allowed us to shuffle allele counts between populations, to find out truly remarkable genetic distances. That way, the permutation procedure unlinks the genetic histories of different populations, which is akin to standardizing the recent effects of gene flow and removing the spatial autocorrelation in genetic drift.

4.2.2 Genetic distance

Genetic distances between time points for a given location were calculated using Rogers' genetic distance (Rogers 1972), which is similar to the Euclidean genetic distance (**Supp. Mat. 4-1**). Rogers' distance makes no assumptions about base-pair substitutions or time since separation and is therefore appropriate to study short-term dynamics influenced by environmental forces. One advantage is that Rogers' distance is simplifiable for biallelic markers and is related to the well-known Euclidean distance (**Supp. Mat. 4-1**, **Supp. Mat. 4-2**). Rogers' distance has been used previously to investigate variation in genetic diversity in several taxa (Bennett & Stone 2019; Pereira et al. 2018). We computed Rogers' distance using the *dist.genpop* function from the *adegenet* R package (see *Software*). While we chose to illustrate TGI using Rogers' distance, our approach and functions may also be used with other true distance metrics (e.g., angular distance/Edward's distance) and measures of differentiation (e.g., F_{ST}), although testing should ensure their relevance (see *Discussion*).

4.2.3 Simulation framework

To simulate population genetic changes through time and test the performance of TGI, we used the spatially-explicit gene flow simulator *CDMetaPOP* (Landguth et al. 2017b). *CDMetaPOP* simulates dispersal and mating of individuals across a landscape and allows the user to define the initial genetic structure, spatial distribution of individuals, dispersal characteristics, and life history traits of the populations. The physical landscape we simulated was a homogeneous, interconnected 5 × 5 square grid with each of the 25 cells representing a population. Each population had a maximum carrying capacity of 50 individuals; the populated landscape therefore contained a maximum of 1,250 (25 × 50) individuals. Distance between populations was set as the Euclidean geographic distance. The genotypic information of each individual consisted of 100 neutral, unlinked, bi-allelic single nucleotide polymorphism (SNP) loci. We included mutation to increase the realism of our simulations, because mutation is a common source of temporal genetic variation. The nuclear mutation rate (per base pair per generation) was set at 10^{-8} (Allio et al. 2017). Although this empirical estimate seems rather low, the total number of simulated loci over the simulation study far exceeds it and we, therefore, expect mutation events in some of our simulations. Mutation, gene flow and drift are all explicitly modelled in the simulator, which uses individual-based modelling. The mutation rate was directly entered in the simulator. The simulated gene flow depends on a large number of parameters related to dispersal ability, but also to reproduction parameters and input ecological distances between population (here equal to Euclidean distances in a homogeneous grid). Genetic drift indirectly depends on many input parameters in the simulation model, notably mortality parameters, and carrying capacity of each cell in the landscape. For readers who might want to run simulations for their systems, genetic drift *lato sensu*, although generally interpreted temporally, may also be strongly influenced by dispersal parameters, because dispersal create a spatial analogue of drift in expanding populations where the study landscape is not yet entirely colonized (Slatkin & Excoffier 2012).

We investigated the influence of a single demographic event on the spatial-temporal apportionment of genetic variation. Specifically, we tested the capacity of our TGI approach to detect demographic and genetic changes due to immigration from an isolated population. This isolated population was separate from our 5 × 5 grid (i.e., population #26). This independent source population shared the same attributes as other populations in our simulated landscape. During simulated demographic events, individuals from population #26 were only allowed to disperse into the 5 × 5 simulation grid.

Using *CDMetaPOP*, we examined the influence of dispersal (movement among our 25 populations) and the spatial extent of a demographic event (number of populations that received immigrants from population #26) on the persistence of spatial genetic legacies. Here, persistence of genetic legacies was defined as our ability to identify them using TGI (see *Section 2.6*). We examined three levels of dispersal capacity (described below) and three different numbers of affected populations (1, 2, or 3) for a total of nine unique scenarios (**Table 4-2**). Each scenario was then simulated 180 times, for a total of 1,620 (9 × 180) unique replicates for this experiment, excluding the control simulations (**Table 4-2**).

For each replicate, we initialized the simulation by randomly distributing alleles among individuals, therefore approaching the maximum (p=q=0.5 maximizes $p \times q$) genetic diversity (Landguth et al. 2017b). Those parameters were chosen as a compromise between realistic allele distributions and computational limitations and were appropriate for producing simulated genetic data that could reasonably recreate the complex evolutionary dynamics in real populations. Each simulation was run for 100 generations before the demographic event was imposed on up to three populations in the landscape. Ten additional generations were simulated after the event. Sampling was performed up to 9 generations before and after the event.

4.2.4 Dispersal

Dispersal was modelled using the weighted geographic distance between populations using a power law function, $10^{-B * distance}$, where B represents the difficulty of dispersal. High values of B correspond to low dispersal capacity. Within a simulation run, B was constant while we randomly picked distances at which individuals dispersed based on the power law function. To do so, we rescaled the values of all distances in the landscape, using the maximum $(4\sqrt{2})$ and the minimum (0) distances possible in this virtual landscape, as described in the *CDMetaPOP* (Landguth et al. 2017b) user's manual (p. 63). This produced values in the [0,1] range. Rescaled values were considered to represent probabilities that an individual disperses to a cell located at that distance (**Table 4-1**). We chose this way of modelling dispersal to allow for both within-population movement and landscape-wide dispersal (Mayrand et al. 2019a).

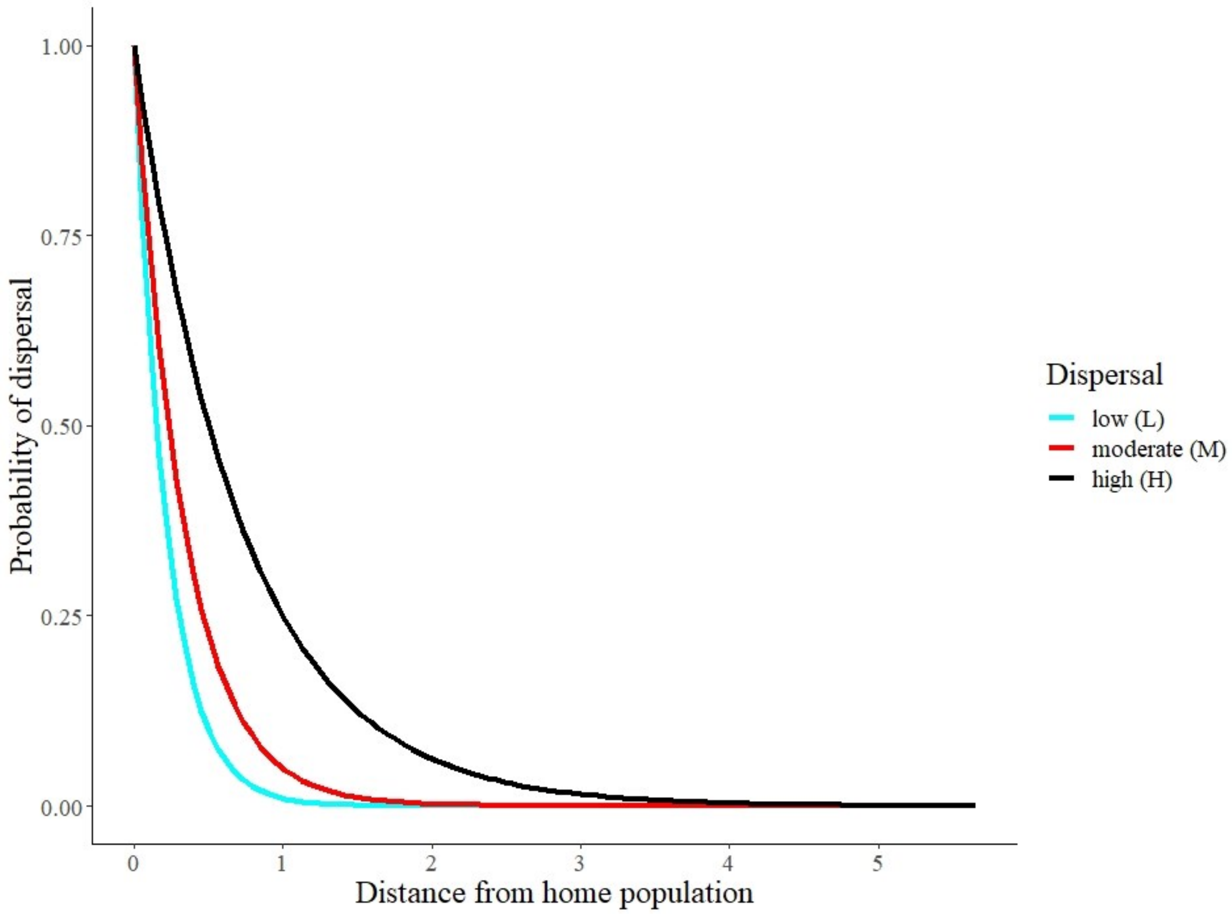


Figure 4-1 : Probability of dispersal of an individual as a function of geographic distance, in three different dispersal scenarios.

The population to which an individual dispersed was selected randomly from the set of populations available at the distance which was randomly picked following the probability distribution described above. We set the landscape boundary conditions to impermeable and individuals, therefore, always stayed within our simulated landscape. Any individual could disperse to any one of the 25 populations at each generation with probability of dispersal decreasing with distance. To investigate the effect of different levels of dispersal, we ran separate simulations using three different values of B: low (B = 2), moderate (B = 1.301) and high (B = 0.6015) dispersal capacity (**Figure 4-1**, **Table 4-2**).

4.2.5 *Spatial extent*

We also evaluated how the spatial extent of the simulated immigration event affected the performance of our TGI method. To do this, we allowed individuals from population #26 to immigrate into one, two, or three populations that were randomly selected from the original 25. We varied the position of where the demographic event occurred in the landscape among simulation replicates because deme topology may influence the outcomes of population genetic analyses (Robledo-Arnuncio & Rousset 2010). For scenarios in which only one population was affected, we partitioned the 180 replicate simulations equally among six populations in the landscape. Because our landscape is square and homogeneously resistant to movement, it is symmetric and there are only six unique positions. Systematically choosing those six positions as a pool of potential targets for our demographic event therefore covered all possible spatial patterns in our simulations. One or more populations among these six populations were randomly selected once and were identical across runs. When multiple (two or three) populations underwent a demographic event, we randomly chose one of these six geographically unique populations and randomly picked one or two additional populations directly adjacent to it. We chose to pick adjacent populations to respect the spatial autocorrelation often exhibited in demographic events. For each of the two- and three-population simulations, we repeated this population selection procedure six times and ran 30 replicate simulations for each set of populations.

4.2.6 *Statistical performance*

We assessed the statistical performance of our TGI testing procedure using the false positive rate (FPR) and false negative rate (FNR). In our study, a false positive was a population that we knew did not undergo the demographic change we imposed, but was found to have done so using the TGI test, whereas a false negative was a population that experienced the demographic event but exhibited no significant change according to the TGI test. The false positive rate is expressed as the ratio of false positives to the total number of negative tests (i.e., true negatives and false positives), and the false negative rate is expressed as the ratio of false negatives to the total number of positive tests (i.e., true positives and false negatives).

A high false positive rate would indicate that our TGI measure often selected the wrong population(s) as having changed significantly and that our testing procedure was less selective. Researchers generally want to minimize the false positive rate when there are, for example, limited resources available for conservation efforts. In contrast, a high false negative rate would mean that we often failed to identify the population(s) that were actually affected, and that our testing procedure had low discriminatory power. Researchers may want to minimize the false negative rate in situations where finding all affected populations is the most important aspect, for example, if there is limited time to take conservation action.

TGI's performance may vary with cut-off values, and a trade-off between power and selectivity is expected to occur. To characterize this compromise, we evaluated the statistical performance of TGI using a range of significance thresholds for calculating a false positive rate and a false negative rate: 0.001, 0.005, 0.01, 0.015, 0.020, 0.025, 0.030, 0.035, 0.040, 0.045, 0.050, 0.055, 0.060, 0.065, 0.070, 0.075, 0.080, 0.085, 0.090, 0.095, and 0.1. We did not conduct corrections for multiple tests (e.g., Bonferroni) because in our simulation framework, use of such a test would simply translate into another range of thresholds. However, choosing to correct for multiple tests is possible in our function. Note that 0.001 is the lowest possible value when using 999 permutations. A compromising threshold value for the trade-off, as evaluated through this question, was used in the analyses focused on time lag.

4.2.7 Time

We sought to assess how the time since the simulated demographic event influences the performance of the TGI method in each of our nine dispersal/spatial extent scenarios (**Table**

4-2). To do so, we calculated the TGI for simulated data collected up to nine generations before and after the event and compared it to the TGI calculated from data collected immediately before or after the event generation. We chose nine generations as the maximum time between samplings (nine generations before the event, or nine generations after the event) because this timeline is longer than most “before/after” population genetic studies in the literature (e.g., Bezault et al. 2011; Kinziger et al. 2015; Moraes et al. 2017) and most long-term ecological research programs monitor at a shorter time interval (e.g., Hobbie et al. 2003; Knapp et al. 2012; Kuemmerlen et al. 2016). Comparisons between TGI results were based on the false positive rate and false negative rate calculated at a significance threshold of $p \leq 0.05$, as this threshold was a good compromise between different performance metrics as indicated by the results to the trade-off analysis described above.

4.2.8 Controls

Control simulations were run in which populations were only affected by local gene flow, drift, and mutation. No demographic events were included. Dispersal was the only parameter that varied among the control simulations, resulting in three control scenarios (**Table 4-2**). We only evaluated the false positive rate of these control scenarios; because there were no true positives or false negatives for populations affected by the demographic event, the false negative rate was always equal to zero. The performance of experimental scenarios was always compared to the control scenario with the same dispersal capacity.

4.2.9 Software

CDMetaPOP runs on *Python 2.7* (Landguth et al. 2017b). We used the *R* software (R Core Team 2019) in the *RStudio* IDE (RStudio Team 2018) for all analyses and illustrations. We used the *adegenet* 2.1.5 (Jombart 2008; Jombart & Ahmed 2011), *pegas* 1.0-1 (Paradis 2010), *poppr* 2.9.3 (Kamvar et al. 2014, 2015) and *adespatial* 0.3-7 (Dray et al. 2019) *R* packages for calculations.

Our *TGI* function is available in the supplementary material (**Supp. Mat. 4-2**) as an *R* script. The function enables the user to choose among five different genetic distance metrics for the calculations of TGI indices via the "*method"* argument.

4.2.10 Applied example: an endangered fish

To demonstrate that our TGI measure provides valuable information about temporal change in a real system with conservation implications, we applied it to real genetic data from a study of a threatened vertebrate, the Northern tidewater goby (Kinziger et al. 2015). We chose this example because it uses a different type of genetic data than we used for our simulations, thus demonstrating that TGI is applicable to a variety of genetic markers. In addition, the study authors suggested that one goby population had undergone more genetic change than the other, more stable local populations, allowing us to test a real hypothesis and go beyond a simple illustration of our method (Kinziger et al. 2015). The dataset was downloaded from DRYAD (doi: 10.5061/dryad.871db). Significance testing of temporal change in these data used 9,999 permutations.

4.3 **Results**

We were able to translate the temporal beta-diversity indices framework to TGI by adapting it to the specific structure of genetic data. Although our results in the present section support the general efficacy of TGI and warrant its use on empirical datasets, the performance of the TGI approach was sensitive to dispersal capacity, spatial extent of the demographic event producing the genetic change, and the time difference between surveys and the demographic event.

Experimental false positive rate values were consistently lower than control false positive rate values, regardless of dispersal parameters (**Figure 4-2**). This suggests that when using TGI, we are less likely to misidentify a significant genetic change in presence of an actual demographic event, than we are in its absence. Among the control simulations, runs with higher dispersal capacity had a lower false positive rate (**Figure 4-2**). Control false positive rate values were generally at least twice as high as the maximum experimental false positive rate values encountered (L1, M1), regardless of the significance threshold used. This means that, even for the lowest-performing scenarios in our simulations, TGI was much more effective at avoiding false positives in the presence of an event than in the absence of one.

4.3.1 Dispersal

Dispersal capacity influenced our ability to detect temporal changes in genetic diversity, as the false negative rate generally increased with dispersal capacity (**Figure 4-3**). However, only one scenario (H3; **Table 4-1**) exhibited false negative rate values above a very conservative limit of 1%, regardless of the *p*-value threshold used (**Figure 4-3**). Of the four scenarios that did not achieve an average false negative rate of 0 (L3, M3, H2, and H3), two involved high dispersal. When we averaged the false negative rate values calculated at the traditional $p \leq 0.05$ threshold across scenarios sharing the same dispersal parameters (e.g., averaging false negative rate value for L1, L2, and L3 grouped together), the mean false negative rates were 0.0037 (0.0007 – 0.0066; 95% confidence interval [CI]) for low dispersal, 0.0049 (0.0015 – 0.0083; 95% CI) for moderate dispersal, and 0.0108 (0.0055 – 0.0161; 95% CI) for high dispersal.

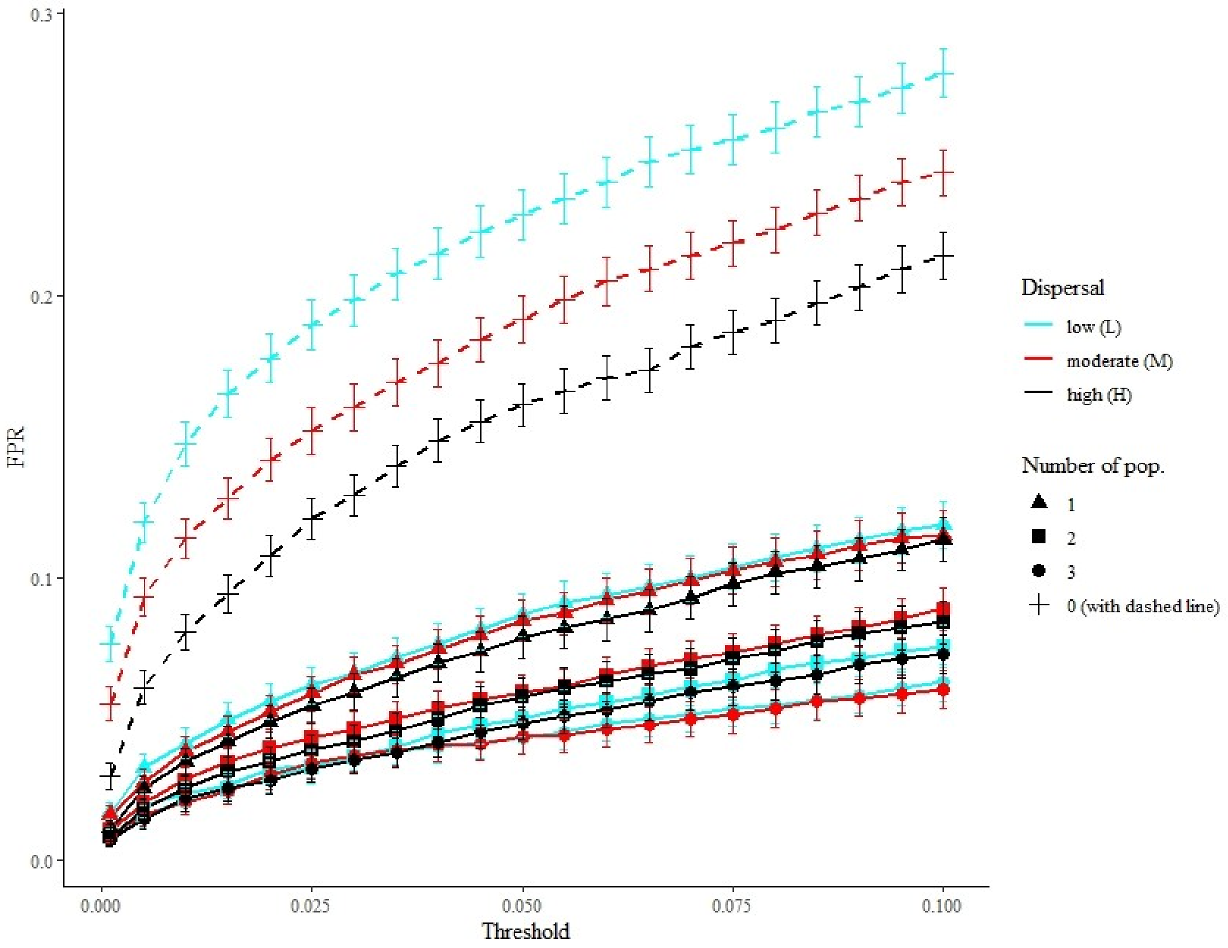


Figure 4-2 : False positive rate across all threshold and scenarios. Control experiments are shown with dashed lines. Those values are for samplings done at the 100 and 101 generations 100 and 101, i.e., right before and after the migration event. 95% confidence intervals of the false positive rate estimates are displayed by vertical bars.

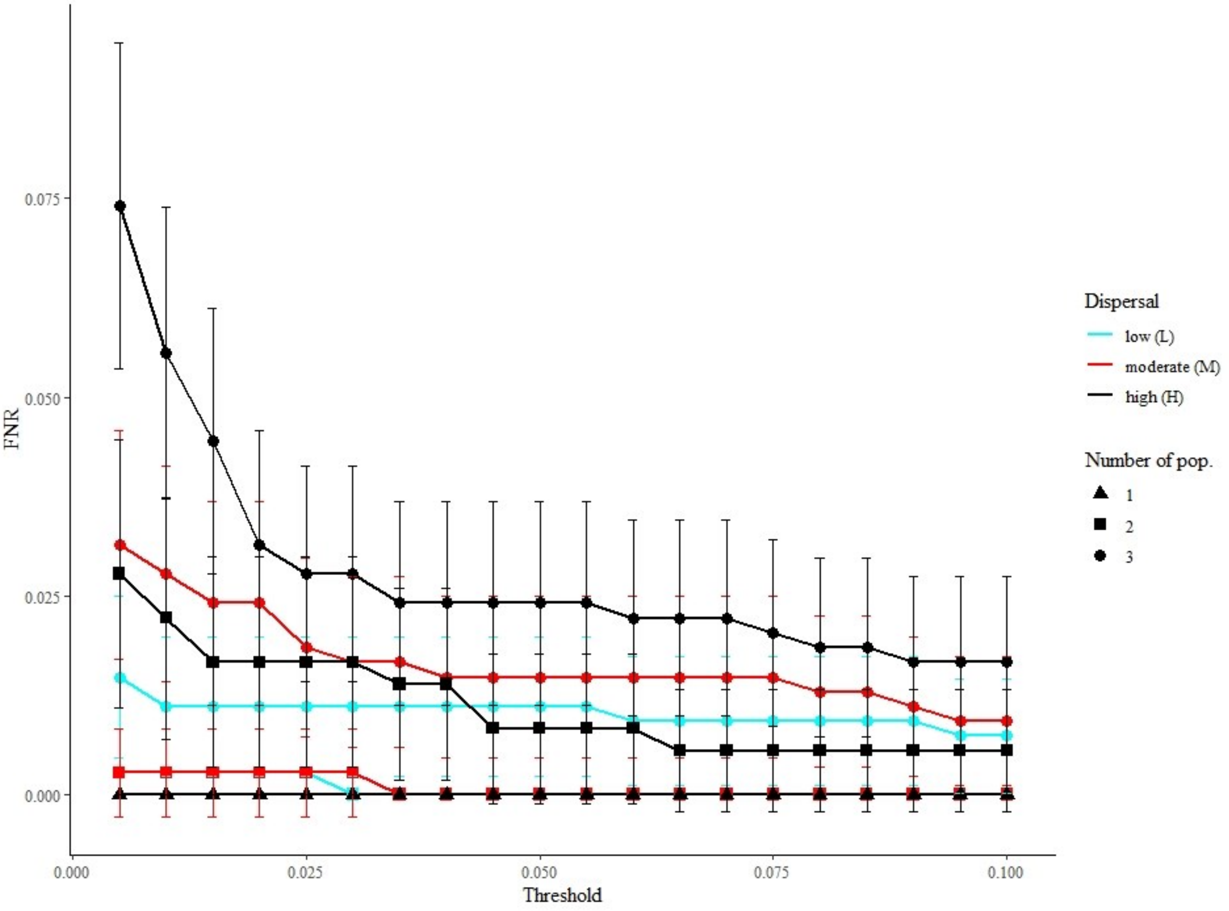


Figure 4-3 : False negative rate across all threshold and scenarios. There are no control experiment results displayed for false negative rate because there are no possible true positives in control experiments, hence no false negatives either. Those values are for samplings done at generations 100 and 101, i.e., right before and after the migration event. 95% confidence intervals of the false negative rate estimates are displayed by vertical bars. Very low thresholds (i.e., < 0.01) are so conservative that they sometimes lead to no population being selected. Symbols overlap for some scenarios (those with only one affected population for example), which reach a null false negative rate.

In contrast, dispersal capacity did not substantially affect the false positive rate (**Figure 4-2**). There were no consistent trends in false positive rate when comparing scenarios with different dispersal capacities but the same number of affected populations: L1 had slightly higher values than M1 and H1; L2 had slightly lower values than M2 and H2; L3 had intermediate values between those of M3 and H3. Average false positive rate values for scenarios sharing the same dispersal parameters, calculated using false positive rates at the $p < 0.05$ threshold as before, were 0.0599 (0.0558 – 0.0641; 95% CI) for low dispersal, 0.0621 (0.0580 – 0.0662; 95% CI) for moderate dispersal, and 0.0600 (0.0562 – 0.0638; 95% CI) for high dispersal (**Figure 4-2**).

4.3.2 Spatial extent

The number of populations affected by a demographic event also influenced our ability to detect meaningful temporal change. Scenarios in which fewer populations were affected exhibited a reduced false negative rate and an increased false positive rate (**Figure 4-2**, **Figure 4-3**). Scenarios in which a single population was affected (i.e., L1, M1, H1) had a perfect false negative rate (0; **Figure 4-3**), while scenarios L2 and M2 only reached this perfect false negative rate at more liberal significance thresholds (i.e., above $p \leq 0.03$; **Figure 4-3**). The mean false negative rates at $p \leq 0.05$, averaged across scenarios sharing the same number of affected populations (e.g., one averaged value for L1, M1, and H1 grouped together), were zero for scenarios with one affected population, 0.0028 (0 – 0.0059; 95% CI) for scenarios with two affected populations, and 0.0167 (0.0105 – 0.0228; 95% CI) for scenarios with three affected populations.

The number of affected populations influenced the false positive rate more than dispersal in our simulations. False positive rate values were consistent across scenarios with different dispersal levels but the same number of affected populations, rather than across scenarios with similar dispersal levels but different numbers of affected populations (**Figure 4-2**, **Figure 4-3**). The average false positive rates from scenarios with the same number of affected populations, determined at the $p \leq 0.05$ significance threshold, were 0.0820 (0.0778 – 0.0863; 95% CI) for scenarios with one affected population, 0.0553 (0.0516 – 0.0591; 95% CI) for scenarios with two affected populations, and 0.0447 (0.0413 – 0.0481; 95% CI) for scenarios with three affected populations.

4.3.3 Time

We found that the genetic signal of the demographic event decayed over time, but that the TGI test was still able to identify significant changes in genetic diversity at a time scale of 1–9 generations. However, as the time interval between pre- and post-event surveys increased, the ability of TGI to detect the demographic event decreased, evidenced by the increase in false positives and false negatives for several demographic scenarios (**Figure 4-4**, **Figure 4-5**). The effect of time between surveys on the sensitivity of TGI was strongly affected by dispersal capacity and the extent of the event.

The timing of sampling prior to a simulated event was, as expected, generally less important than the timing of the post-event survey. The decrease in the genetic signal over time — which would be found with any comparative method, not just TGI — was considerably strong in our simulations. For example, if the second (post-event) survey was taken nine generations after the first (pre-event) survey, we observed high false negative rate values that approached 75–90% in high- and moderate-dispersal scenarios (**Figure 4-4**). The false negative rate also increased with the time lag in low-dispersal scenarios, but the increase was more linear, and values never reached 30%, even after nine generations (**Figure 4-4**). One interesting result was that the number of affected populations was the main factor driving increasing false negative rate values with the age of the pre-event survey (3>2>1; left side of (**Figure 4-4**), while dispersal capacity was the main factor driving increasing false negative rate values when the time between the event post-event survey increased (H>M>L; right side of **Figure 4-5**). For scenarios with the same number of affected populations, moderate-dispersal scenarios showed the worst performance with pre-event survey time lags, whereas high-dispersal scenarios generally showed the worst performance with post-event survey time lags (**Figure 4-4**). Over our nine-generation sampling window, the false negative rate changed the least for the L1 scenario and the most for the H3 scenario (**Figure 4-5**).

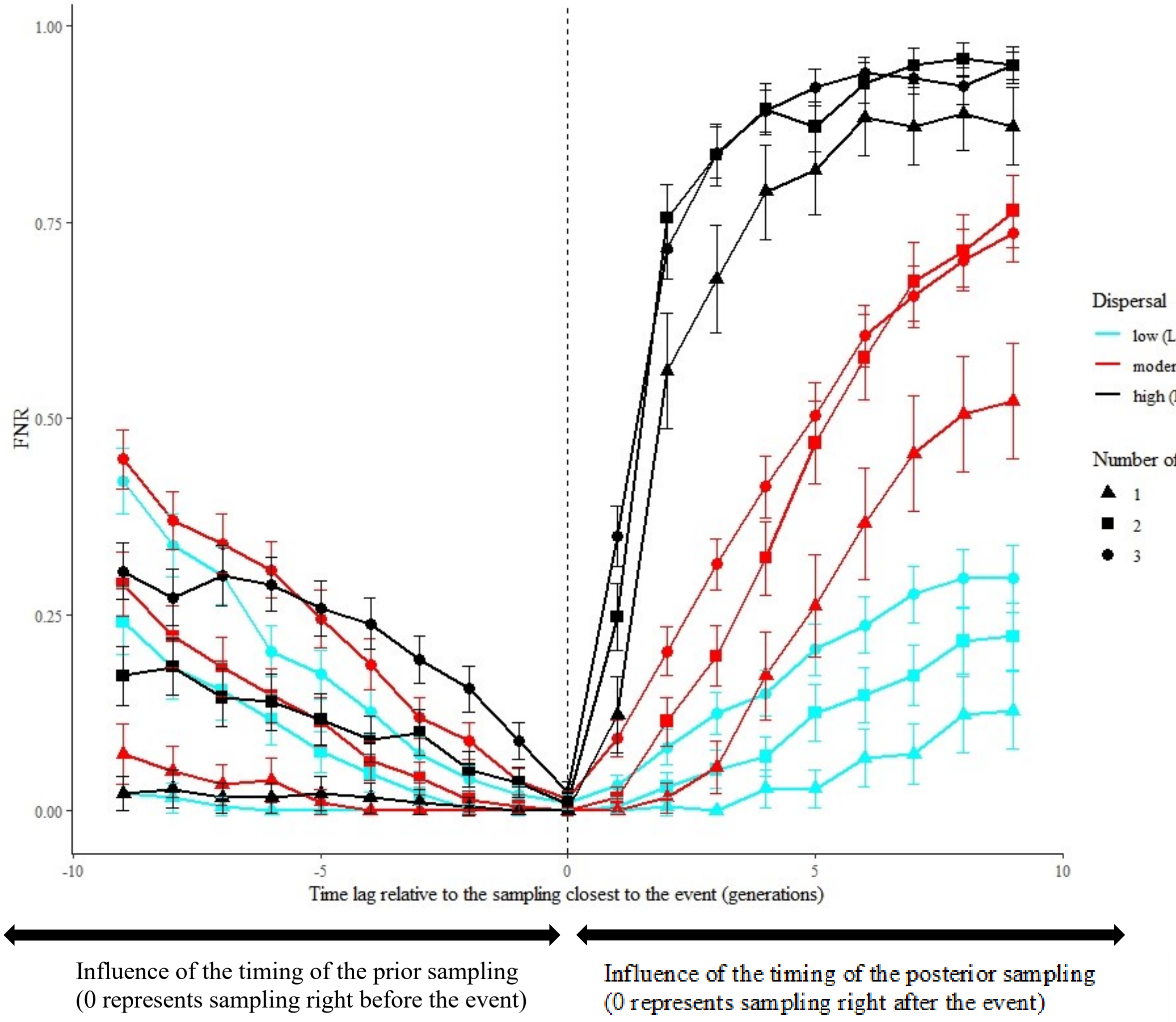


Figure 4-4 : False negative rate from TGI tests performed between surveys carried out up to 9 generations before or after the migration event (arrow) when compared with surveys done the generation after the event for prior sampling, or the generation before the event for posterior sampling. 95% confidence intervals are displayed by bars.

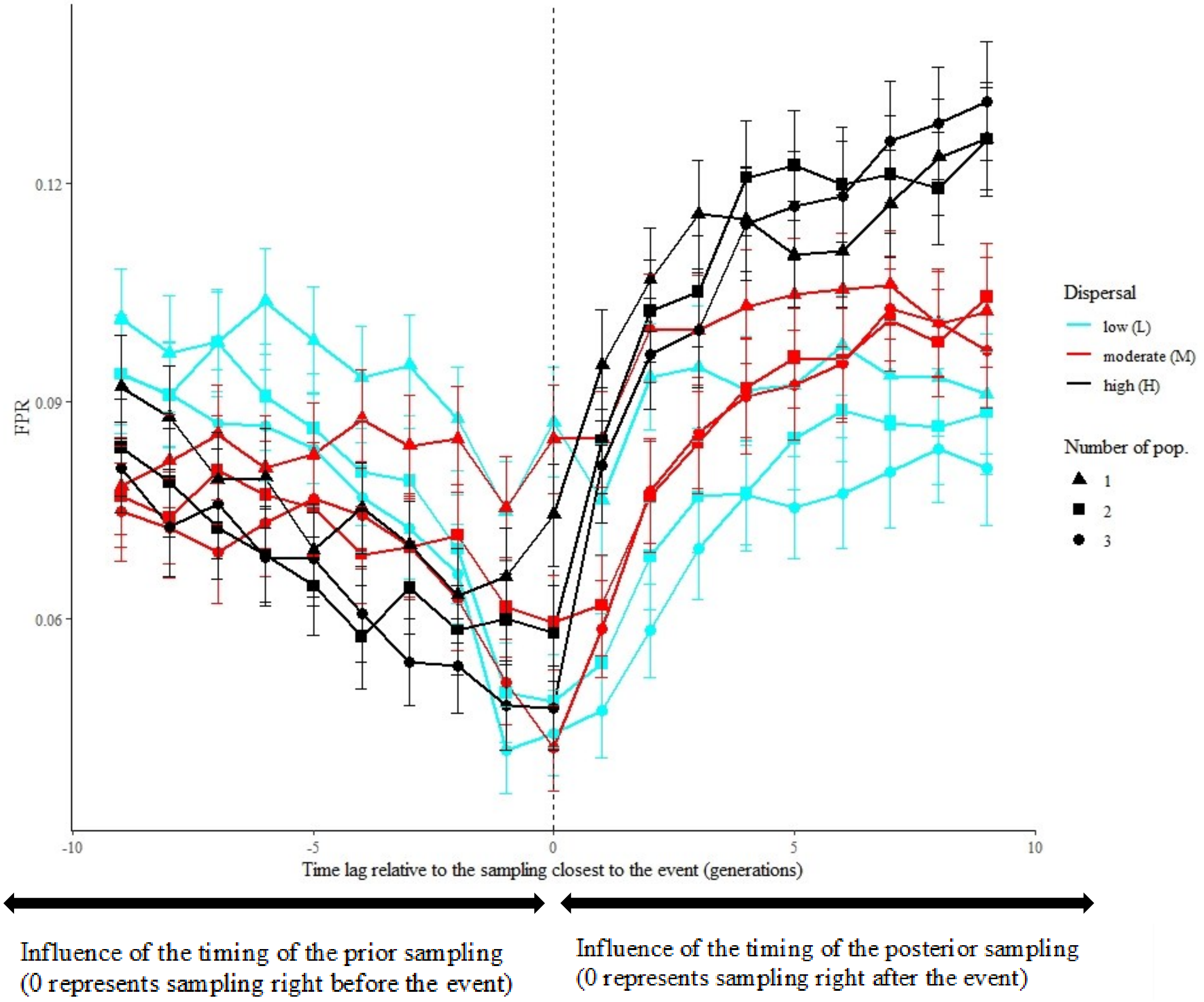


Figure 4-5 : False positive rate from TGI tests performed between surveys executed up to 9 generations before or after the event (arrow) when compared with surveys done the generation after the event for prior sampling, or the generation before the event for posterior sampling. 95% confidence intervals are displayed by bars.

While the relative differences in false positive rate performance given different time lags were not as high as for false negative rate, false positive rate nonetheless increased with the survey time lag. There were no clear patterns for whether dispersal or the number of affected populations most influenced the change in false positive rate associated with pre-event sampling time (**Figure 4-5**); however, dispersal was the main factor driving false positive rate for time gaps associated with post-event sampling (**Figure 4-5**). The strong relationship that we observed between false positive rate and the number of populations affected by the demographic event therefore became less pronounced as dispersal became more influential. As with the false negative rate, the false positive rate did not change much for the L1 scenario and changed the most dramatically for the H3 scenario (**Figure 4-5**). These differences in how time affects our two most extreme scenarios are a useful consideration for potential TGI users.

The simulation that was most likely to preserve the signal of the demographic event was the low-dispersal scenario with a single affected population (L1). In this scenario, the TGI approach was still able to keep false negatives below 15% and false positives below 10%, even when the second survey was done nine generations after the event (**Figure 4-4**, **Figure 4-5**) and regardless of whether the first or second survey was responsible for the time lag with the event.

4.3.4 Thresholds

A trade-off based on significance threshold values between false negative rate and false positive rate was present across scenarios. False negative rate values decreased with the chosen significance threshold, with a sharp decrease (most notable for H3) before 0.025 followed by a slower decrease until 0.1. False positive rate values increased with the chosen significance threshold, with a sharp increase at low thresholds followed by a continued but saturating increase until $p \leq 0.10$.

4.3.5 Applied example

The Northern tidewater goby (*Eucyclogobius newberryi*) is a small, endangered fish that lives in brackish estuaries and lagoons along the coast of California. This species represents an interesting model for population genetic studies because dispersal between suitable habitat patches only occurs during rare, discrete events. A previous study investigated extinction–colonization dynamics in the tidewater goby by evaluating genetic diversity across the landscape at several points in time (Kinziger et al. 2015). These authors suggested that the Elk River goby population had experienced unexpected temporal genetic change between 2006 and 2011 (Kinziger et al. 2015). We used the TGI method to re-analyze these data and determine if significant temporal genetic changes had indeed occurred in any population in this landscape.

Using our TGI measure, we found that the genetic structure of the Elk River population (**Supp. Mat. 4-3**) of Northern tidewater goby (Kinziger et al. 2015) had indeed changed significantly relative to the other populations surveyed in the study area (permutation *p*-value = 0.0004), even after using strict *p*-value adjustments (Holm-Bonferroni adjusted permutation *p*-value = 0.0032). The average expected heterozygosity decreased by 0.046 in the Elk River population from an original value of 0.2646, which represents a loss of around 17%. Using TGI, we were able to quantify the qualitative findings of the previous study that there was a loss of genetic diversity in the Elk River population.

4.4 Discussion

In this study, we investigated how dispersal, the spatial extent of a demographic event, and the timing of sampling affected our ability to identify populations that have experienced significant changes in genetic diversity using a novel statistical tool: TGI. The factors we chose to investigate are directly relevant to conservation studies. Dispersal is a key element in understanding population connectivity (Kool et al. 2013; McRae 2006). The spatial extent of significant demographic changes is relevant to examine as recent studies have advocated for a more comprehensive integration of space in evolutionary ecology research (Battey et al. 2020; Bradburd & Ralph 2019; Velázquez et al. 2016). The timing of sampling is key, first to make sure we capture the effects of a potential disturbance, and second, to better grasp the decay in its legacy signal.

4.4.1 TGI: a new and useful framework

TGI provides a novel and robust framework for testing whether observed changes in genetic diversity through time are significant relative to variation associated with genetic drift, local gene flow, and mutation. Our successful application of temporal beta-diversity indices to genetic data involved translating a site-by-species approach to a site-by-genotype approach and changing the permutation algorithm to accommodate the specific structure of various genetic data formats such as SNPs in our simulations and microsatellites in our application. In addition to describing our new framework, we also evaluated its power and specificity and found that TGI is functional over a wide range of parameter values. One main contrast between our new TGI approach and previous investigations of the performance of temporal beta-diversity indices (Legendre 2019; Winegardner et al. 2017), which was developed for community composition data, is that we also examined how the timing of surveys, and its interaction with demographic parameters, may affect the downstream conclusions. Our results indicate that TGI consistently and accurately identifies populations that have experienced a demographic event.

4.4.2 Dispersal and spatial extent

Detection of temporal genetic changes was sensitive to dispersal; false negatives increased with dispersal capacity, although false positives did not show a clear trend (**Figure 4-2**, **Figure 4-3**). The influence of dispersal on the false negative rate was also affected by the time lag between an event and the subsequent sampling effort; the effects of different dispersal capacities were evident even when surveys were separated by only one generation (i.e., samples were collected immediately before and after the event) and were magnified as the time between surveys increased. The effects of sampling time and dispersal capacity on the false negative rate suggest that species with high dispersal capacity in well-connected landscapes, such as many forest pests (e.g., Larroque et al. 2019; Wittische et al. 2019), might require more frequent sampling to overcome the negative effect of gene flow on our ability to correctly identify affected populations.

The spatial extent of a demographic event increased our ability to correctly identify populations that have not truly changed (lower false positive rate), but it also decreased our ability to correctly identify populations that did change (higher false negative rate). The magnitude of this trade-off varied with dispersal capacity. Although a broader spatial extent may help researchers detect an event, as the chance of sampling an affected population increases, it may also increase the risk of not identifying the genetic legacy of the event at all, especially in high-dispersal landscapes. It is less effective for analyzing gradual, landscape-wide disturbances. In addition, when multiple populations were affected in our simulations, we always chose to affect adjacent populations; we did not investigate whether lowering the degree of spatial autocorrelation (Dale & Fortin 2014; Legendre & Legendre 2012) in the spatial genetic legacy (e.g., two independent catastrophic events, pollution and a flood for example, affecting the landscape) influenced our ability to detect the events.

4.4.3 Time between surveys

As expected, spatial genetic legacies decayed over time, mostly due to gene flow and drift. Specifically, in this study we found that TGI was suitable for identifying changes over 1-9 generations (e.g., years) depending on landscape and demographic parameters. Two main points emerged from our analysis of how the timing of sampling affected the detection of significant genetic changes. First, when comparing an old survey to a survey realized soon after a demographic event, the spatial extent of the disturbance affected the power of TGI, with smaller spatial extents preserving high power even with large time gaps. Second, when comparing a survey realized immediately before a disturbance to one collected several years after, dispersal was the most important factor driving the performance of TGI, with low-dispersal scenarios better preserving the performance of TGI in the context of decay brought by background processes through time. High-dispersal systems could lead to as many as 10% of false positives even when sampling only a few years after an event. This result has serious implications: arbitrary and potentially inappropriate significance thresholds may result in misallocation of resources to monitoring or treating unaffected populations while missing some affected populations. In contrast, by considering the population dynamics and, if possible, planning relatively simple and short model-specific simulations, one can enhance the usefulness of TGI and proceed with a more appropriate sampling/monitoring strategy. Given the fact that false negative rate reach high values at the highest time gaps for some scenarios (**Figure 4-4**), we believe that our choice of a maximum of 9 generations (e.g., years) between surveys was appropriate.

4.4.4 Empirical application

We successfully applied TGI to an empirical dataset from an endangered fish, the Northern tidewater goby, for which temporal genetic change had been described but not quantitatively tested (Kinziger et al. 2015). The authors of the original publication hypothesized that one goby population had undergone meaningful genetic change relative to the rest of the landscape; our application of TGI supported this hypothesis. We therefore clearly showed that the straightforward TGI testing procedure can be used to strengthen the results from temporal genetic studies that use repeated surveys.

4.4.5 Considerations about the use of TGI

Different empirical datasets and research objectives may require TGI users to customize our procedure, but the TGI function is transparent and flexible (Culina et al. 2020), and different permutation and genetic distance algorithms could easily be used by simply changing a few characters or lines of R code in the annotated TGI function provided in the supplementary material (**Supp. Mat. 4-2**). Evaluating the performance of distance metrics with different statistical properties (Legendre & De Cáceres 2013) could be an avenue for future work. TGI can also readily be used on other types of genetic data, such as microsatellites. TGI provides a robust statistical framework more trustworthy than arbitrary comparison of pairwise genetic dissimilarities, or node-based genetic diversity values.

Despite these advantages, there are still several important considerations for the effective use of TGI tests. The implementation of TGI in new systems will ultimately be more successful if researchers have an a priori understanding of the population dynamics of their system and the nature and scale of possible disturbances in their study area. Indeed, this prior knowledge could guide researchers in the choice of survey intervals after important historical demographic events (Anderson et al. 2010; Fenderson et al. 2019). False negative rate and false positive rate values ultimately represent trade-offs in potential conservation costs (Moilanen et al. 2009; Welch et al. 2020), and it is therefore essential that researchers grasp their importance and choose these values deliberately. Stricter (lower) values for the TGI *p*-value threshold expectedly result in a lower false positive rate but may also result in a higher false negative rate (lower power). Identifying the most sensible threshold for a chosen objective would be valuable to better understand the trade-offs of different sampling schemes in specific empirical systems. Purpose-designed spatially explicit simulations can be used to address this challenge (Epperson et al. 2010; Haller & Messer 2019; Landguth et al. 2017b). In some cases, it may be desirable to minimize false negatives relative to false positives – thus ensuring that we detect all the affected populations no matter the cost of detecting, and therefore monitoring and preserving, some populations that do not need preservation.

Neutral micro-evolutionary processes at the metapopulation-level such as gene flow, genetic drift and mutation may vary a lot between species and demographic contexts (e.g., central vs. edge population), which may limit the feasibility of temporal change detection for some natural systems. For example, we investigated the effect of gene flow across three levels from low to high and our results showed that the higher the gene flow the faster time would erase the signal of a past demographic event. We chose to focus on dispersal because it is known to affect other detection tools focusing on micro-evolutionary processes (Lotterhos & Whitlock 2014; Mayrand et al. 2019b). Another potential limit brought by neutral processes such as gene flow, is the often unknown influence from unsampled subpopulations which may sometimes alter the results (Beerli 2004; Koen et al. 2013; Naujokaitis-Lewis et al. 2013; Slatkin 2005) but also, counter-intuitively, help population genetics analyses (Shirk et al. 2021). However, beyond dispersal ability, there are other intrinsic characteristics of populations which affect our ability to detect change.

Genetic drift is a result of the ubiquitous and relentless battle between evolutionary and ecological processes which may limit genetic variation in populations, and hence every population genetic analysis. When genetic drift is very strong, besides being a cause of concern for conservation (Ellstrand & Elam 1993; Hedrick & Kalinowski 2000; Willi et al. 2013) because it creates low levels of genetic variation, it can also erase the signal of past demographic events which makes it even harder to monitor already threatened small populations. Even when population sizes are large, they may not be exempt of genetic drift which creates deviations from an ideal population, with each deviation lowering the effective population size (N_e). Although we have not done it in this study, it would be important to test the effect of strong genetic drift on the performance of TGI with a varying effective population size, as we expect performance to decrease with lower effective population size. Empirical applications of TGI could be affected by low effective size. Tools to measure N_e from genetic data are available but generally require additional genetic information beyond a genotypic matrix such as linkage disequilibrium (Barbato et al. 2015), a huge dataset (Barker 2011)) or at least two temporal datasets at least four generations apart (Tallmon et al. 2010; Waples 2005). We did not simulate scenarios where linkage disequilibrium is theoretically possible because, because as stated in the methods, we chose to simulate physically independent loci. Our simulations were parameterized in a way to keep effective population size close to the census population with a very narrowly changing sex-ratio and variation in reproductive success, a well-connected metapopulation, and little variation in population sizes across the landscape for many generations. We chose to do this to keep genetic drift from varying too much between scenarios, because genetic drift does interact with dispersal, to avoid genetic drift overwhelming the effect of the other factors we focused on in this study. We simulated genetically healthy populations which may lower the usefulness of our simulations for some conservation applications where populations have already been suffering from the consequences of heavy genetic drift. However, this paragraph should be taken as yet another reason to conduct system-specific simulations for each study where genetic drift is assumed to be a major player. Given sufficient knowledge about a species and the local demographic context, simulating genetic drift is possible in CDMetaPOP.

TGI was not developed as an alternative for inferring demographic history from large genetic datasets collected at a single time (e.g., Leblois et al. 2014). Instead, it was designed to help research teams collecting repeated surveys from non-model organisms with limited genotypic information (Draheim et al. 2018; Kinziger et al. 2015; Moraes et al. 2017), and teams wanting to compare new surveys to older ones. Nonetheless, more studies in different types of genetic systems, involving different historical demographic events, are needed to explore how the performance of TGI varies with factors that were not tested in our simulations, including 1) the chosen genetic distance algorithm; 2) spatiotemporal autocorrelation in genetic legacies; 3) effective population size; and 4) spatial heterogeneity in landscape resistance to movement.

Using a different null model approach could lead to different performance and insight, and permutation is but one of many possible ways to create reference distributions which can be used to build a test by comparison with the observed value. Although comparing different null model approaches was beyond the scope of this paper, preliminary work for this study showed that some permutation algorithms were generally less adequate for false positive rate and false negative rate, such as was the case for temporal beta-diversity indices (Legendre 2019). Future research may notably include evaluating more complex, multilevel, permutation approaches, and exploring null models which explicitly preserve some components of genetic structure such as isolation-by-environment (Wang & Bradburd 2014). Previous spatiotemporal research in ecology has shown that null models with increasing levels of complexity provided different insight (James et al. 2010), but that the usefulness of complex null models could be offset very long computation time (Leblois et al. 2014), and by the risk of overfitting (Bohl et al. 2019). The performance of more complex null models may also be more sensitive to threshold choice (Merckx et al. 2011), which is another reason why evaluating a range of threshold values, such as in this study, is pertinent.

Temporal beta-diversity indices have inspired TGI but there are some reciprocal opportunities brought by insights generated through it's porting to genetic data. We showed that dispersal, through gene flow, had profound effects on our ability to detect change. Dispersal is also a key element of community dynamics (Hubbell 2011), and we know that it influences community assembly and mediates community response to an environmental change (Catano et al. 2017; Condit et al. 2002; Conradi et al. 2017; Qian 2009). Dispersion limitation is not constant across different species and preserving the range of dispersal abilities through a change in the permutation algorithm may increase temporal beta-diversity indices performance, especially for fine grain/large extent landscapes, and for communities with strong distance decay in similarity. Another insight brought by our study, which also interacts with dispersal, is the window of the time during which the signal of a community-changing event may be detectable using temporal beta-diversity indices. This is especially important to consider as species having different dispersal abilities may differently recolonize and the gains and losses part of a TGI analysis (Legendre 2019) may therefore be dependent on the time since the event. Integrating TGI and temporal beta-diversity indices in a multi-scale community genetics (James et al. 2011a) change approach, and testing the trajectory (De Cáceres et al. 2019; Sturbois et al. 2021) of the genetic diversity or community composition across more than two dates, represent two promising avenues of research.

4.4.6 Conclusions

At the crux of many conservation biology questions, identifying changes in genetic diversity, beyond the expected changes due to background processes, can help researchers identify locations or populations that have experienced important past demographic events. These events could be detrimental (e.g., loss of diversity or maladaptation) or beneficial (e.g., higher effective population size or genetic rescue), and are often relevant for conservation efforts. Such locations and populations could then be prioritized for increased monitoring and further investigation into the origin of these changes. As shown in our application of TGI to empirical data from the endangered Northern tidewater goby, our method provides a framework for detecting and pinpointing exceptional temporal genetic changes. Our approach to detect temporal genetic differentiation does not require extensive genomic information and can therefore be used to explore the temporal dynamics of genetic diversity changes using relatively small genetic datasets (e.g., hundreds of SNPs). We believe that the TGI approach is a promising tool for the spatiotemporal analysis of wild, non-model organisms for which extensive genomic resources are yet to be developed.

TGI and its future developments will, therefore, be of primary value to better manage the landscape in the context of global change and stronger, and more frequent, perturbations. Monitoring the genetic health of already at-risk populations (Draheim et al. 2018; Kinziger et al. 2015; Moraes et al. 2017) is but one of the uses TGI can offer. Climate change will displace populations which are differently adapted to warmer and drier environments (Masson-Delmotte et al. 2021) and being able to detect this genetic change is of primary value to better manage species affected by climate change. Given that the relative importance of tolerance, migration, and adaptation, may vary in different parts of a landscape under climate change (Sork et al. 2010), TGI could be used as a sentinel tool to assess whether populations have changed substantially more than the general response of this species to climate change in that landscape. TGI could also be used to detect a new contact between two previously distinct genetic clusters in the landscape, which could alert managers that a shift is happening (Jay et al. 2012; Pérez-

Portela et al. 2019). TGI could also be used to study the dynamics and the synchrony of outbreaks (Larroque et al. 2019), by, for example, detecting a large migration event from a previously isolated genetic cluster, which may confirm the high connectivity synonymous with serious outbreaks. Biological invasions may be slowed, or fail, because of genetic factors such as successive reduction of genetic diversity following introduction, or the spatial analogue of genetic drift at the leading edges of the expanding range (Austerlitz et al. 2000; Schrey et al. 2014; Slatkin & Excoffier 2012). A local gain or a loss of genetic diversity, as identified by TGI, could predict the evolutionary potential of a population which may affect the further spread of that invasive species (Lawson Handley et al. 2011), because of genetic load, inbreeding depression, and drift load (Charlesworth & Willis 2009; Szucs et al. 2014; Willi et al. 2013). Similarly, TGI could be used to detect illegal human-mediated translocations of non-native individuals (Dufresnes et al. 2016; Frantz et al. 2006, 2017) which can be an issue because of maladaptation, outbreeding depression, and pathogen introduction in naïve populations.

Acknowledgements

This research was supported by a grant to PMAJ and the TRIA Network from the Natural Sciences and Engineering Research Council of Canada (grant no. NET GP 434810-12), with contributions from Alberta Agriculture and Forestry, fRI Research, Manitoba Conservation and Water Stewardship, Canadian Forest Service (Natural Resources Canada), Northwest Territories Environment and Natural Resources, Ontario Ministry of Natural Resources and Forestry, Saskatchewan Ministry of Environment, West Fraser, and Weyerhaeuser. This research was also supported by a Discovery Grant to PMAJ. JW was also supported by a scholarship from the Forest Complexity Modelling (FCM) NSERC CREATE. Computations were made on the supercomputer CEDAR managed by Compute Canada (www.computecanada.ca). We thank Jeremy Larroque, Hinatea Ariey and Charlotte Van Engeland for their comments on an earlier version of the manuscript.

DATA ACCESSIBILITY

All simulation data used for this paper will be deposited online upon acceptance. Functions used to analyze the simulations will be available on a public repository on *GitHub*. TGI, the function that would be most useful to potential users of our approach, will continue to be maintained and developed and may be contributed to a CRAN package in the near future.

AUTHOR CONTRIBUTIONS

J.W. designed the study, created the simulation inputs, ran the simulations, transformed the temporal beta-diversity indices function to TGI, and performed the analyses. P.L. and P.M.A.J. provided advice on the study design, analysis, and the visualization. J.W., P.L. and P.M.A.J. wrote the paper.

Supplementary material

Supp. Mat. 4-1 : Rogers' genetic distance

Given L the number of loci, $a_k$the number of alleles observed in locus k, and $p_{pop_1,j,k}$ and $p_{pop_2,j,k}$ the proportions of allele *j* at locus *k* in populations 1 and 2, respectively:

$$D_{Roger}(pop_1, pop_2) = \frac{1}{L}\sum_{k=1}^{L}\sqrt{\frac{1}{2}\sum_{j=1}^{a_k}\left(p_{pop_1,j,k} - p_{pop_2,j,k}\right)^2}$$

The sum of the allele proportions for any locus in any population is 1.

For biallelic markers such as SNP, the calculation becomes even simpler as seen in the example in **Table 4-1**, A. Indeed, because the proportion of an allele is always 1 minus the proportion of the other allele, the sum per locus is always twice the term for the a single allele and the equation simplifies to:

$$D_{Roger}(pop_1, pop_2) = \frac{1}{L}\sum_{k=1}^{L}\left|p_{pop_1,1,k} - p_{pop_2,2,k}\right|$$

Supp. Mat. 4-2 : TGI function and output.

```
# mat1: the genotypic matrix associated with the first sampling; must be a
genind object
# mat2: the genotypic matrix associated with the second sampling; must be a
genind object
# nperm: the number of permutations used in the evaluation of significance
# seed.: you may specify a seed by using this argument

# method : a number between 1 and 5. Five genetic distances are available in
function dist.genpop # of the adegenet package.
# They are : (1) Nei’s D, (2) Edwards’ angular D, (3) Reynolds’ coancestry
coefficient, (4)
# Rogers’ D, (5) Prevosti’s absolute genetic D. Methods 2, 3 and 4 produce
Euclidean distances,
# whereas methods 1 and 5 produce non-Euclidean distances, which produce
negative eigenvalues and # complex eigenvectors in principal coordinate
analysis.

# correc: correction for multiple # inference; see ?p.adjust
# thresh_for_GL: indicate here the threshold you want to use

TGI <—function (mat1, mat2, nperm = 999, replace = FALSE, seed. = NULL,
method = 4, correc = "holm", thresh_for_GL = 0.05) {

  #### genind to genpop objects

  mat1p <- genind2genpop(mat1)
  mat1p <- mat1p[,order(colnames(mat1p@tab))]
  mat2p <- genind2genpop(mat2)
  mat2p <- mat2p[,order(colnames(mat2p@tab))]

  ##### Function to compute genetic distances

  dissim <- function(mat1p, mat2p, method) {
    dis <- vector(mode = "numeric", length = nrow(mat1p@tab))
    for (i in 1:nrow(mat1p@tab)){
      if (i == 1){
        trick <- 2
      } else {
        trick <- 1
      }
      temp_genpop <- mat1p
      temp_genpop@tab[trick,] <- mat2p@tab[i,]
      dis[i] <- dist.genpop(temp_genpop[c(trick, i),], method = method)
    }
    list(dis = dis)
  }

  ##### Initialization of seed, tolerance

  if (!is.null(seed.)){
```

```
  set.seed(seed.)
}

epsilon <- sqrt(.Machine$double.eps)

##### Dimensions check

n <- nrow(mat1p@tab)
p <- ncol(mat1p@tab)
if ((nrow(mat2p@tab) != n) | (ncol(mat2p@tab) != p)){
  stop("The matrices are not of the same size!")
}

##### Empirical genetic distances
tmp <- dissim(mat1p, mat2p, method)
dis.ref <- tmp$dis

##### Permutations
# We obviously need more than 0 permutations
if (nperm > 0) {
  # Stores seed values later used to randomize permutations
  my.vec <- sample(1:(10 * nperm), size = nperm)
  # Init. how many times the permutation distances are higher than observed
  outlier.count = rep(1, n)
  # This loop computes the permutations and the permutation p-values

  for (iperm in 1:nperm) {
    # Set seed
    set.seed(my.vec[iperm])
    # Observed genotypic matrix for Time 1 (earliest sample)
    mat1.perm <- mat1p
    # Permutation genotypic matrix for Time 1 (earliest sample)
    mat1.perm <- shufflepop(mat1.perm, method=4)

    set.seed(my.vec[iperm])
    # Observed genotypic matrix for Time 2 (latest sample)
    mat2.perm <- mat2p
    # Permutation genotypic matrix for Time 2 (latest sample)
    mat2.perm <- shufflepop(mat2.perm, method=4)
    # Computation of the genetic distance between Time 1 and Time 2
    tmp <- dissim(mat1.perm, mat2.perm, method)

    dis.perm <- tmp$dis
    # How oftem are the permutation distances higher than the observed?
    ge <- which(dis.perm + epsilon >= dis.ref)
    if (length(ge) > 0) {
      outlier.count[ge] <- outlier.count[ge] + 1
    }
  }
    # Raw permutation p-value computation
  p.dist <- outlier.count/(nperm + 1)
}
```

```
  p.adj <- p.adjust(p.dist, method = correc)

  ##### Gain or loss?

  n.pop1 <- seppop(mat1)
  n.pop2 <- seppop(mat2)

  mean.hexp1 <- do.call("c", lapply(n.pop1, function(x)
mean(summary(x)$Hexp)))
  mean.hexp2 <- do.call("c", lapply(n.pop2, function(x)
mean(summary(x)$Hexp)))

  mean.hexp1[is.nan(mean.hexp1)] <- NA
  mean.hexp2[is.nan(mean.hexp2)] <- NA

  simple_diff <- mean.hexp2 - mean.hexp1

  output <- list(TBI = dis.ref, p.TBI = p.dist, p.adj = p.adj, gainloss =
simple_diff[p.adj < thresh_for_GL])

  class(output) <- "TGI"

  return(output)
}
#############################################################################
> goby_test <- TGI(goby_first, goby_second, nperm = 9999, method = 4)

> goby_test

$TBI #index values

[1] 0.06432131 0.06089485 0.15212258 0.02258920 0.07247326 0.04463856
0.06672004 0.02238467

$p.TBI #unadjusted permutation p-values

[1] 0.4283 0.4891 0.0004 0.9943 0.3188 0.7756 0.3947 0.9949

$p.adj #adjusted p-values

[1] 1.0000 1.0000 0.0032 1.0000 1.0000 1.0000 1.0000 1.0000

$gainloss #difference in expected heterozygosity

        ELK

-0.04567755

attr(,"class")

[1] "TGI"
```

Supp. Mat. 4-3 : Satellite map of the Californian sampling stations used in the goby analysis. The red circle marks the Elk population which was the only one found to have significantly changed.

5. CONCLUSION

5.1 Synthèse

Les connaissances issues de l'écologie spatiale sont maintenant reconnues comme cruciales en biologie de la conservation, et les concepts spatiaux sont régulièrement mis au premier plan dans les solutions proposées aux problèmes environnementaux. Contribuer à mieux comprendre comment les espèces répondent à l'hétérogénéité spatiale, notamment celle issue de la multiplicité de perturbations, m'a conduit à examiner une large gamme de questions tout en suivant cependant le même axe de recherche. Mon projet a en effet suivi un but général et double : mieux comprendre la dynamique des populations dans des écosystèmes faisant face à des perturbations sans précédent, tout en contribuant à la pratique de l'écologie du paysage grâce à l'application et au développement de méthodes efficaces et innovantes. Pour être plus spécifique, il s'agissait tout d'abord d'étudier les impacts d'une mosaïque de perturbations sur la distribution locale d'une communauté forestière, et ce en prenant en compte l'autocorrélation spatiale et en préservant la communauté dans des analyses basées sur les nœuds (Chapitre 2). Ensuite, il s'agissait de quantifier par une analyse basée sur les liens, les contraintes imposées par le paysage sur le mouvement d'une espèce forestière éruptive et dévastatrice, et ceci malgré l'obstacle présenté par l'absence de données (i.e., opinion d'experts) liée à la nouveauté de l'habitat atteint par l'espèce (Chapitre 3). Finalement, je me suis concentré sur la description et l'évaluation d'une nouvelle méthode robuste face à la paucité des données, qui permet de détecter un réel changement temporel de diversité génétique calculé pour chaque nœud tout en considérant l'entièreté du paysage (Chapitre 4). En essayant de contribuer aux deux facettes de ce but général à travers ces trois chapitres d'analyse, j'ai tenté d'ajouter une pierre à l'édifice visant à mieux comprendre le vivant, notamment à des fins de protection des écosystèmes et de leurs services. Le Chapitre 2 supporte et complémente la littérature existante portant sur les effets de perturbations liées à l'exploitation de la forêt boréale sur les mammifères, apportant les bénéfices d'une approche multi-espèces et multi-échelles. Ce chapitre décrit la manière envahissante dont un grand nombre de perturbations affectent toutes les espèces du paysage. L’analyse la plus novatrice dans ce chapitre est sans doute celle qui compare explicitement les effets de deux types de perturbations anthropiques, et qui montre qu’elles expliquent des parts importantes, mais complètement différentes de l’utilisation de l’habitat par les mammifères boréaux. Le Chapitre 3 s’inscrit dans un large effort visant à mieux comprendre l’épidémie de dendroctone du pin ponderosa (James & Huber 2019). Ce chapitre a souligné le rôle important du paysage, notamment du climat, dans la propagation du ravageur. Un des aspects les plus innovants de ce chapitre est l’immense étendue de paramètres qui a été considérée dans la calibration des modèles, avec plusieurs millions de modèles comparés durant le processus d’optimisation. Le Chapitre 4 vient enrichir l’arsenal méthodologique de l’écologue en fournissant une nouvelle méthode de suivi temporel de la génétique des populations. Cette nouvelle méthode remplit une lacune grâce à sa capacité à accepter deux jeux de données appariés, et à ses faibles prérequis en ce qui concerne les données. Un des aspects innovants de cette méthode est qu’elle est inspirée d’une méthode issue de l’écologie des communautés, encourageant plus d’échanges technologiques fructueux entre ces disciplines.

L’objectif du Chapitre 2 était de déterminer quels éléments, naturels et anthropiques, du paysage affectaient le plus une communauté de mammifères de la forêt boréale de l’Ouest canadien (Wittische et al. 2021). En particulier, ce chapitre a permis de comparer les effets des facteurs naturels et anthropiques, ainsi que de comparer différents types d’altérations du paysage entre eux, de manière innovante. Nos hypothèses de départ pour ce chapitre concernaient l’utilisation des altérations anthropiques linéaires du paysage par les prédateurs, et l’influence des proies à forte biomasse, et ont été supportées et étayées. Cependant, et contre toute attente, j’ai montré que la densité d’habitats altérés liés à l’exploitation pétrolière des sables bitumineux et à la foresterie avait un effet plus fort que l’habitat naturel sur la distribution des prédateurs. Une autre surprise fut que parmi les perturbations anthropiques, les perturbations non linéaires (en bloc, par exemple, les coupes à blanc ou les zones d’extraction à ciel ouvert) expliquaient autant la densité des proies que les perturbations linéaires, en dépit d’une littérature se concentrant récemment sur ces dernières. Les deux types de perturbations expliquaient la variation spatiale de l’utilisation du paysage par les espèces de manière complètement différente et complémentaire, tant pour les prédateurs que pour les proies. Ceci souligne l'importance de systématiquement considérer les perturbations linéaires et non linéaires et de les comparer, dans des approches basées sur les nœuds.

Le Chapitre 2 offre une nouvelle perspective quant aux pressions anthropiques les plus importantes sur la distribution et l'activité locale d'un large groupe de vertébrés de la forêt boréale. Comprendre ces pressions est crucial, car elles avantagent certains groupes et désavantagent d'autres. En effet, certaines populations de ces espèces, comme le cerf de Virginie et le coyote, sont en expansion (Fisher et al. 2020; Hody & Kays 2018; Laliberte & Ripple 2004), alors que d'autres comme celles du caribou forestier sont vulnérables (Palm et al. 2020), entre autres à cause de ces nouvelles expansions (Côté et al. 2004; Frenette et al. 2020). De plus, ces espèces vulnérables ont besoin d'habitats spécifiques (Courbin et al. 2009) et de la mise en place de multiples mesures de protection pour avoir une chance de persister (Winder et al. 2020). Il faut également garder à l'esprit que le changement de composition de la communauté est à la fois une conséquence et un moteur de la perte de biodiversité et que les interactions entre espèces peuvent changer avec le paysage. Il est donc d'autant plus important de mieux comprendre la communauté dans son ensemble, à travers différentes approches méthodologiques complémentaires. Ceci est impératif étant donné que les changements d'origine anthropique vont continuer à se développer, notamment dans les régions à sables bitumineux (Rosa et al. 2017), et interagir avec le changement climatique (Musetta-Lambert et al. 2019; Yeung et al. 2019), et donc, continuer à créer des gagnants et des perdants dans de nombreux écosystèmes.

La combinaison de l'urgence de conserver la biodiversité, avec la disponibilité des données spatiales et le développement méthodologique, a stimulé l'intégration de plusieurs espèces dans les questions concernant les perturbations anthropiques. Au-delà des nouvelles connaissances à propos de la forêt boréale, les résultats du Chapitre 2 soulignent également l'importance de modéliser la communauté complète plutôt que d'utiliser plusieurs modèles monospécifiques, de comparer des groupes thématiques de prédicteurs, et d'incorporer l'espace explicitement. En effet, l’approche multispécifique est plus réaliste, prenant en compte les interactions entre de nombreuses espèces sans avoir à donner un poids arbitraire à chacune, tout en préservant notamment les caractéristiques des espèces rares (Baselga & Araújo 2009; Hui et al. 2013; Maguire et al. 2016). Mon utilisation novatrice du partitionnement de la variance s’est montrée particulièrement adaptée et efficace pour traiter ce genre de problématique. Plus spécifiquement, l’incorporation de variables spatiales dans les différents modèles, et l’analyse de la fraction de variation expliquée partagée entre ces variables spatiales et les variables environnementales ont permis de mieux appréhender l’autocorrélation spatiale présente dans le paysage, et les échelles pertinentes (Legendre & Legendre 2012; Legendre et al. 2012; Peres-Neto & Legendre 2010).

Dans le Chapitre 2, j’ai couvert de nombreuses sources de perturbation dans une large zone d’étude de la forêt boréale de l’ouest du Canada, mais il existe d’autres processus écologiques clés à considérer afin de mieux comprendre cet écosystème. Parmi les autres perturbations majeures de la forêt boréale figurent notamment la sécheresse, les feux de forêt, et les épidémies d’insectes (Boucher et al. 2018). Ces trois perturbations affectent profondément la composition forestière, qui à son tour influence de nombreuses espèces animales telles que celles étudiées dans le Chapitre 2. Parmi ces trois perturbations, les épidémies d’insectes sont les plus mal comprises malgré leur étendue spatiale (p. ex., $>10^6$ ha pour plusieurs pestes forestières au Canada). Le Chapitre 3 se concentra donc sur un des insectes envahissants qui détruit le plus d’arbres au Canada et dont les épidémies interagissent avec la sécheresse et les feux de forêt : le dendroctone du pin ponderosa (DPP). Ce coléoptère s’attaque à des conifères commercialement exploités, et affecte des millions d’hectares de forêt en Amérique du Nord. Plusieurs projets majeurs supportés par le gouvernement fédéral et les provinces, ainsi que par les industriels de la forêt ont donc été menés pour ultimement pouvoir atténuer les épidémies graves de DPP et leurs conséquences néfastes sur l’économie et la biodiversité (James & Huber 2019). Afin de déterminer quels facteurs contribuent à, ou ralentissent, la propagation du DPP, il est nécessaire d’utiliser une approche différente de celle utilisée dans le Chapitre 2 (Lowe & Allendorf 2010). L’utilisation de variables spatiales et de larges rayons dans l’analyse réalisée dans le Chapitre 2, ainsi qu’un index capturant en partie les comportements de mouvements et d’activité des espèces, peut nous donner une idée des mouvements au voisinage des sites d'échantillonnage. Cependant, avec cette approche, les effets de l'habitat naturel et des perturbations sur le mouvement ne sont pas facilement distinguables des effets sur l'abondance locale (Broadley et al. 2019; Neilson et al. 2018a; Stewart et al. 2018). De plus, le type de mouvement qui pourrait être inféré en utilisant cette approche et ces données ne peut pas aisément être lié à de la dispersion effective qui implique à la fois le mouvement de dispersion ainsi qu'une reproduction fructueuse sur le site de dispersion (Cayuela et al. 2018; Clobert et al. 2009). Or, les informations sur la dispersion effective entre sous-populations sont cruciales pour mieux comprendre et prédire la propagation du DDP et les conséquences des perturbations liées aux épidémies.

L'objectif du Chapitre 3 fut donc d'explorer une perturbation d'origine entomologique sans précédent, en examinant les éléments du paysage qui facilitent et contraignent le mouvement du DPP dans les forêts de conifères de l'Ouest canadien (Wittische et al. 2019). En particulier, ce chapitre a permis à la fois de quantifier la résistance au mouvement créée par différentes variables environnementales agissant sur le mouvement du DPP (Bonte et al. 2012; McRae 2006), ainsi que de trouver la combinaison de variables la plus performante pour prédire la connectivité génétique, de manière innovante. Nos hypothèses de départ concernant la relation entre environnement et résistance, bien que qualitatives et limitées par le manque d'opinion d'experts, furent toutes supportées par l'analyse réalisée en complète indépendance vis-à-vis des hypothèses. La meilleure surface de résistance sélectionnée dans notre étude comprend une combinaison d'altitude et de climat favorable. L'extrapolation de cette surface composite pour explorer la connectivité future lors d'une expansion vers l'Est présentait une connectivité moins forte et moins homogène que celle utilisant une surface basée sur la densité en hôtes. Malgré cela, même en utilisant cette surface combinant altitude et climat, nos résultats ne suggèrent pas une barrière, ni même une forte résistance, à une expansion future de l'épidémie de DPP vers l'Est, voire le Nord. Les informations apportées par le Chapitre 3 sont cruciales pour comprendre quelles régions sont les plus susceptibles à une expansion rapide du DPP.

Les résultats du Chapitre 3 soulignent la puissance d'une approche d'optimisation par algorithme évolutionniste, notamment lorsque l'opinion des experts manque. En effet il est très complexe de transformer des valeurs environnementales brutes présentes dans chaque cellule du paysage (conceptualisé en grille; Fig. 3), en valeur de résistance au mouvement (Peterman 2018; Peterman et al. 2014). Allant au-delà de calculer les paramètres d'une simple relation linéaire comme cela est traditionnellement fait dans la littérature, cette approche permet de transformer les valeurs en coûts en utilisant une multitude de fonctions. Ceci permet notamment de découvrir des relations fonctionnelles cryptiques entre environnement et mouvement. Le type de modèle utilisé par le logiciel *ResistanceGA* 3.0.0 (Peterman 2018) est également à la pointe de la recherche en génétique du paysage, car il permet de résoudre explicitement le problème de non-indépendance des valeurs de connectivité calculées dans un paysage (Clarke et al. 2002). Il est nécessaire que cette propriété soit dorénavant respectée de manière systématique dans les études utilisant des analyses basées sur les liens, car au-delà d'une raison conceptuelle, ces modèles ont récemment été testés et sont plus performants que les approches antérieures telles que le test de Mantel ou les régressions sur matrices de distances classiques (Shirk et al. 2018). De plus, cette approche supporte un processus de sélection de modèles robuste (Row et al. 2017).

Optimiser la relation environnement-coût requiert souvent d'optimiser simultanément plusieurs surfaces de résistance basées sur différentes variables environnementales, mais l'optimisation monosurface présente également un intérêt (Peterman 2018). L'optimisation multisurface confère un avantage conceptuel important, car différentes surfaces n'influencent pas la connectivité fonctionnelle de manière indépendante. Optimiser simultanément la calibration de résistances basées sur plusieurs variables, comme je l'ai fait pour toutes les surfaces composites (Wittische et al. 2019), a été supporté comme étant conceptuellement et statistiquement supérieur à la combinaison post-optimisation de plusieurs surfaces de résistances calibrées indépendamment (Peterman & Pope 2021). Il devient de plus en plus évident que la compréhension de la relation entre des prédicteurs environnementaux et la connectivité fonctionnelle est un objectif très différent de trouver la meilleure surface (souvent composite) pour prédire la connectivité. En effet, alors qu'il est extrêmement complexe d'interpréter les résultats de l'optimisation d'une surface composite, cette surface représente sans aucun doute une meilleure base pour prédire la connectivité et sans doute pour l'extrapoler à un paysage plus grand ou différent. Nous encourageons donc les futures études à définir leur objectif parmi ces deux alternatives dans leur démarche de modélisation de la connectivité, que ce soit entre individus ou entre populations (Balkenhol et al. 2015).

Les Chapitres 2 et 3 ont exploré deux différentes facettes de la dynamique des populations dans des systèmes subissant des perturbations sans précédent, mais il y a une dimension importante qu'ils n'ont pas explicitement explorée : ces chapitres n'abordent pas la question spatio-temporelle des changements de biodiversité. En effet, la distribution géographique des espèces, leurs patrons de mouvement, et le paysage dans lequel les populations changent au cours du temps. La dimension temporelle, bien que hardie à modéliser, est souvent indispensable pour aborder les problématiques en écologie du paysage (Borcard et al. 2018; De Cáceres et al. 2019; Dornelas et al. 2013; Ontiveros et al. 2021), et notamment en génétique du paysage (Bradburd & Ralph 2019) où les changements de structure génétique spatiale reflètent les changements dans le paysage (Draheim et al. 2018). Il est devenu évident qu'il fallait aller au-delà des « instantanés » d'interprétation pour mieux cerner de nombreuses problématiques en écologie et biologie de la conservation (De Cáceres et al. 2019; Epps et al. 2013; Fenderson et al. 2019). Il est également important de différencier les influences historiques des influences contemporaines, par exemple sur les processus microévolutionnaires (Anderson et al. 2010; Holzhauer et al. 2006; Pavlacky et al. 2009). Les perturbations anthropiques et le changement climatique sont très souvent mis en avant comme facteurs clés pour expliquer les changements biologiques tels que la perte de diversité. Cependant, notre compréhension de leurs effets est limitée par le manque de méthodes qui comparent explicitement la dynamique de populations séparées d'une ou plusieurs générations.

L'objectif du Chapitre 4 fut donc de combler une partie de ce manque méthodologique pour pouvoir mieux répondre aux questions portées sur les changements biologiques à travers le temps, qui vont au-delà de questions spatiales posées dans les deux précédents chapitres. La nouvelle méthode décrite et testée dans le Chapitre 4, TGI, représente une nouvelle option pour analyser certains jeux de données afin de détecter les changements temporels atypiques de diversité génétique. Cette nouvelle méthode, basée sur un test par permutations, permet en effet d'évaluer si la composition génétique d'une sous-population change au cours du temps au-delà de ce qu'on pourrait attendre en ne considérant que les stochasticités démographiques et génétiques. La méthode TGI fonctionne pour un grand espace de paramètres démographiques. Il est aussi possible de déterminer les compromis entre faux-négatifs et faux-positifs pour un système en particulier, en réalisant un petit nombre de simulations. De plus, l'approche ne se base sur aucune hypothèse liée à la coalescence, et son utilisation est techniquement abordable pour un grand nombre d'utilisateurs. L'approche TGI est également facilement personnalisable et nous encourageons fortement de futurs utilisateurs à contribuer à, par exemple, de nouveaux types de distance génétique et des fonctionnalités.

La nouveauté de l'approche TGI réside dans plusieurs aspects. Elle a été conçue spécialement pour les échantillons répétés, c'est-à-dire prélevés dans les mêmes sites à au moins deux dates. Ceci est notamment réalisé pour le suivi de certaines espèces d'intérêt par leur vulnérabilité (Kinziger et al. 2015; Moraes et al. 2017) ou leur caractère envahissant et dévastateur, comme le DPP et la tordeuse du bourgeon de l'épinette (*Choristoneura fumiferana*). Ces deux insectes ont et feront l'objet d'échantillonnages répétés (James & Huber 2019; Johns et al. 2019) et il est prévu d'utiliser TGI sur des séries temporelles de jeux de données pour mieux comprendre la dynamique génétique des populations à différentes phases épidémiques. Par exemple, comment la diversité génétique de ces populations subissant des fluctuations démographiques fortes et cycliques change-t-elle lors d'un déclin démographique ? Pour mieux comprendre les origines du changement de diversité génétique, il est tout à fait possible de régresser des indices de changements du paysage sur les indices calculés par TGI (ou seulement ceux qui sont significatifs dépendamment de l'objectif), ou encore de les utiliser comme des poids dans d'autres approches de modélisation.

La propriété transgénérationnelle de nombreuses données génétiques peut poser des problèmes lorsqu'il s'agit d'interpréter les changements temporels dans la dynamique des populations, ce qui est l'objectif de nombreuses approches telles que TGI. Notre capacité à détecter des changements est liée aux caractéristiques démographiques de l'espèce étudiée telles que la taille efficace de la population, le temps de génération, ou la fréquence des épidémies, ainsi qu'à l'hétérogénéité du paysage. Dans le Chapitre 4, j'ai simulé des populations changeant très rapidement et le signal d'un changement démographique disparaissait rapidement. Dans d'autres contextes démographiques, l'échelle temporelle peut être plus complexe à intégrer à cause d'un décalage entre changements de variation génétique neutre ou adaptative, et les changements environnementaux (Anderson et al. 2010). Il peut par exemple y avoir un délai dans la détection de patrons permettant par exemple la découverte d'un frein à la dispersion (Landguth et al. 2010; Murphy et al. 2008). D'autres problèmes associés à ce décalage concernent la détection des loci sous sélection (Lotterhos & Whitlock 2014; Mayrand et al. 2019a), dont la contribution à un changement phénotypique peut être remplacée par celle d'autres allèles avec le temps ce qui complique encore davantage la détection des associations gènes-environnement (Yeaman 2015). L'évaluation du risque de maladaptation (Capblancq et al. 2020) peut similairement être faussée par un délai dans la perception des changements génétiques et démographiques (Rellstab et al. 2021). Finalement, un autre problème de taille dans l'identification d'un changement temporel grâce aux données génétiques est qu'il peut y avoir un biais vers certaines interprétations lié à l'échantillonnage (Bailleul et al. 2018) ou la statistique utilisée (Landguth et al. 2010). Il peut également y avoir inadéquation entre les résultats de ces analyses et ceux obtenus par des méthodes démographiques (Howeth et al. 2008; Moore et al. 2017; Yu et al. 2010). Les simulations que je propose dans le Chapitre43 peuvent permettre d'identifier et de comprendre une partie de ces différences.

Malgré la diversification continue des questions en écologie spatiale, dont cette thèse ne représente qu'une petite partie, et la possible divergence dans les concepts et les méthodes qui pourrait s'en suivre, une convergence conceptuelle et une synergie dans l'inférence sont possibles grâce à la similarité des données et des questions. Ces bénéfices pourraient devenir un catalyseur pour les échanges interdisciplinaires.

5.2 Transferts conceptuels et méthodologiques entre approches génétiques et démographiques

La méthode TGI étant une adaptation d'une méthode utilisée en écologie des communautés pour la génétique du paysage, il est raisonnable de penser qu'un tel transfert interdisciplinaire concernant d'autres méthodes d'écologie des communautés puisse compléter l'arsenal en génétique du paysage et vice versa. C'est par exemple le cas des contributions locales à la diversité bêta (LCBD en anglais) qui pourraient être utilisées sur des données génétiques pour détecter quelles sous-populations sont génétiquement uniques relativement au reste des sous-populations du paysage. Cela pourrait aider à répondre à la question suivante : « Quelles sous-populations diffèrent le plus relativement à la diversité génétique régionale ? ». Ces indices d'unicité peuvent ensuite, tout comme les TGI, être associés à diverses informations sur le paysage et la dynamique de la métapopulation. Les LCBD n'ont pas encore été appliquées à des jeux de données génétiques, mais leur interprétation et leur performance pourraient être similaires à celles de leur application actuelle. Une autre technique qui provient de l'arsenal de l'écologie des communautés et qui est en lien avec la question temporelle explorée avec les TGI est l'analyse de trajectoire des communautés (CTA en anglais; De Cáceres et al. 2019, Sturbois et al. 2021). En effet, les CTA relient différents « instantanés » de la communauté, qui peuvent être décrits par des coordonnées dans un espace multivarié. Ce dernier décrit la communauté par des combinaisons linéaires et orthogonales d'abondances ou d'utilisations de l'habitat par plusieurs espèces, comme réalisé dans le Chapitre 2. Les propriétés de ces liens, ou segments, et de leurs relations avec les segments précédents ou subséquents, peuvent apporter de nombreuses informations sur les changements de la communauté (De Cáceres et al. 2019; Sturbois et al. 2021). Un transfert de ces CTA vers la génétique pourrait être utile. En effet, les groupements (clusters) basés sur les informations génétiques de populations ou individus sont typiquement réalisés de manière indépendante entre plusieurs périodes d’échantillonnage, alors que garder l’espace multivarié (Jombart et al. 2010) défini lors de la création des clusters génétiques originels peut permettre de mieux saisir les changements de structure génétique dans le paysage (Fang, Wittische, James 2022; en préparation). Des transferts méthodologiques dans l’autre sens, de la génétique du paysage, vers l’écologie des communautés pourraient également améliorer certaines analyses. Par exemple, les modèles MLPE précédemment mentionnés comme étant plus performants que les régressions habituelles en génétique du paysage, pourraient être utilisés sur des mesures de différenciation des communautés.

Pour continuer l’échange méthodologique entre les disciplines associées à l’écologie du paysage, et donc à faire avancer ces sciences, il est nécessaire d’identifier les points en commun et les différences parmi les questions et les données des approches génétiques et de communauté. Par exemple, les formats des données sont généralement différents. En effet, les données génétiques sont souvent résumées sous forme de proportions pour le calcul. Ces proportions sont ensuite souvent elles-mêmes résumées sous forme de distances dans les analyses basées sur les liens, sous forme de diversité pour les analyses basées sur les nœuds et la variation génétique neutre, et sous forme de combinaisons linéaires pour les groupements (Holderegger et al. 2006). Les comptes d’allèles par locus (0, 1 ou 2 pour les diploïdes) sont souvent utilisés notamment pour des analyses basées sur les nœuds et sur la variation génétique adaptative afin de détecter les loci sous sélection (Rellstab et al. 2015, Caye et al. 2016, Privé et al. 2020) et de détecter les loci liés à certains traits (genome-wide association studies ou GWAS; Privé et al. 2019). Les données de présence-absence communes dans l’étude simultanée des distributions de plusieurs espèces (Wilkinson et al. 2019), sont semblables aux données phénotypiques qui sont utilisées dans les GWAS (Hardy & Singleton 2009; Hirschhorn & Daly 2005; Korte & Ashley 2013). Bien que les GWAS ou les modèles de génomique du paysage soient flexibles et, par exemple, également utilisés pour l’analyse de l’expression génique (Jumentier et al. 2020), ils ne sont pas identiques à ceux utilisés dans les données démographiques. Une des raisons derrière ce peu de chevauchement méthodologique pourrait être que dans le cas des données d’abondance, les maxima attendus a priori ne sont pas les mêmes d'une espèce à l'autre alors qu'ils le sont systématiquement entre les loci pour les données génétiques (sans compter les transcriptomes).

Un autre point commun entre les différentes disciplines liées à l'écologie spatiale, et notamment entre les trois chapitres, est qu'ils utilisent la plupart du temps des mesures indirectes des phénomènes démographiques, avec leurs avantages et leurs inconvénients. Dans le Chapitre 2, nous avons pu étudier les effets des perturbations dans une large zone d'étude pendant près de trois ans grâce à l'utilisation de l'habitat par des pièges photographiques. Les pièges photos ont permis de collecter des informations à propos de nombreuses espèces simultanément (Burton et al. 2015; Steenweg et al. 2017), ce qui n'aurait pas été faisable en utilisant d'autres types de mesures (p. ex., télémétrie, pièges à ADN, étude de traces, etc.). Les approches par échantillonnages répétés d'ADN environnemental pourraient être prometteuses, mais ne se prêtent cependant pas encore facilement aux études de génétiques des populations (Adams et al. 2019). Un avantage certain des mesures indirectes utilisées dans les Chapitres 3 et 4 est en revanche qu'elles peuvent permettre plus facilement d'étudier la dispersion effective, et aussi de détecter les événements ponctuels de migration, qu'avec des méthodes démographiques (Cayuela et al. 2018). La combinaison de différentes mesures indirectes démographiques et génétiques dans les études a le double bénéfice d'encourager l'interdisciplinarité, et de se complémenter pour couvrir une plus grande étendue spatio-temporelle, ce qui améliore la précision de l'interprétation des phénomènes écologiques (Aguillon et al. 2017; Almany et al. 2017; Cayuela et al. 2018; Lowe et al. 2017; McCauley et al. 2015; Moore et al. 2017).

Un autre point commun pertinent entre données génétiques et démographiques réside dans le fait que les variants rares, tout comme les espèces rares, peuvent avoir des rôles importants relativement à leur fréquence, qu'il s'agisse d'un rôle biologique et/ou simplement, car ils concentrent de l'information utile à l'analyse. Les espèces de macroorganismes rares sont souvent mises en avant à cause de leur vulnérabilité (Vincent et al. 2020) ou de leur importance écosystémique disproportionnée (Leitão et al. 2016; Mouillot et al. 2013; Soliveres et al. 2016). Bien que les espèces rares de microorganismes aient également en moyenne une plus forte importance relativement à leur densité, elles sont trop peu étudiées (Jousset et al. 2017; Lynch & Neufeld 2015). Il est commun de filtrer les données génétiques pour exclure des variants présents à de très petites proportions (Linck & Battey 2019), tout comme il est commun de filtrer les espèces rares pour les modèles de distributions jointes (Clark et al. 2017). Cependant, certains loci, pourtant neutres relativement à la sélection naturelle, et très rares, peuvent être utilisés pour inférer l'histoire démographique (Eldon et al. 2015; Schraiber & Akey 2015), la structure à échelle fine (O'Connor et al. 2015), ou encore l'origine géographique d'une population (Cubry et al. 2017). L'inférence synchronisée des jeux de données génétiques et démographiques permet, dans le cas d'une convergence, d'atteindre une certitude concernant les estimations des processus étudiés (Legrand et al. 2021; Wang & Shaffer 2017; Watts et al. 2007), ou dans le cas d'une divergence, de mieux cerner les contributions de différents groupes d'individus au patron observé (Lowe & Allendorf 2010; Reid et al. 2016; Yu et al. 2010).

Une des plus grandes différences entre les jeux de données génomiques et ceux de compositions de communautés est que comparée à une communauté où chaque espèce joue un rôle écologique, la majeure partie du génome n'a pas de rôle biologique direct et évident. Cela dit, ces parties non codantes présentent une source pour une potentielle activité fonctionnelle dans le futur (Biémont & Vieira 2006), elles peuvent être conservées sans doute à cause de leur influence sur l'expression génique (Burgess & Freeling 2014; Polychronopoulos et al. 2017), et le rythme de leur évolution peut être associé au développement de certains traits (Rubin et al. 2019). Il y a une forte hétérogénéité de fonction dans le génome qui est associée à une influence en moyenne faible de chaque locus, et des interactions complexes entre les différentes parties du génome. Ces trois éléments peuvent rapprocher conceptuellement ce réseau génomique des réseaux d'interactions spécifiques : ce sont tous deux des réseaux multiplexes (Halu et al. 2019; Kinsley et al. 2020; Peng et al. 2021; Pilosof et al. 2017). Brièvement, les réseaux multiplexes sont une série de réseaux interconnectés dont chaque couche possède les mêmes nœuds (p. ex., gènes, individus, populations), mais dont les liens entre nœuds intracouches sont de natures différentes d'une couche à l'autre. Explorer les réseaux d'interactions spécifiques dans le temps requiert des outils sophistiqués (Ovaskainen et al. 2017a,b). Les avancées importantes dans la théorie des réseaux multiplexes sont donc prometteuses pour les applications en génétique et en écologie des communautés (Carpi et al. 2019; Wang et al. 2021).

Finalement, la mise au point de meilleurs modèles nuls est une prometteuse avenue de recherche qui unit les différentes disciplines associées à l'écologie du paysage. Comme nous l'avons réalisé dans le Chapitre 4, tester la signification d'un patron spatio-temporel requiert généralement la comparaison, directe ou indirecte (Veech 2012), d'une statistique mesurée sur un échantillon empirique, à une distribution de référence. Cette distribution de référence est constituée de valeurs virtuelles de la statistique provenant de randomisations et/ou de modélisation de processus autres que celui ayant généré le patron d'intérêt, et elle peut être complexe à produire (Gotelli & Ulrich 2012). Les modèles nuls ont été largement utilisés en écologie, par exemple, pour tester les associations entre espèces (Gotelli 2001) ou entre loci (Agapow & Burt 2001). Le développement de modèles nuls reste un sujet de recherche important, car avec l'augmentation de leur complexité, il est devenu plus aisé de distinguer la signification des patrons d'intérêts (James et al. 2010). Par exemple, une meilleure prise en compte de l'autocorrélation spatiale dans les modèles nuls a permis l'utilisation de méthodes de randomisation sophistiquées dans des études présentant des plans d'échantillonnage non réguliers (Wagner & Dray 2015).

5.3 Priorités de recherche et directions futures en écologie du paysage

Certains développements récents en biostatistique ont été incorporés dans cette thèse, mais il convient de mentionner d'autres développements de recherche biostatistique à suivre de près pour répondre encore plus précisément à diverses questions d'écologie du paysage. Les transformations des données dans les modèles à réponses multivariées, tels que ceux utilisés dans le Chapitre 2, sont souvent très utiles (Legendre & Gallagher 2001; Legendre & Legendre 2012), mais elles ne sont pas toujours les plus performantes lorsqu'il faut distinguer les effets environnementaux sur la moyenne, des effets environnementaux sur la variance (Warton et al. 2012). La modélisation explicite de la relation entre moyenne et variance peut donc être une solution dans les cas les plus complexes (Miranda-Soberanis & Yee 2019; Niku et al. 2019; Wang et al. 2012; Warton et al. 2012, 2016; Yee 2015, 2020, 2021). Cependant, les approches linéaires telles que la RDA ont des performances similaires dans certaines conditions (Carlos-Júnior et al. 2020; Ives 2015; Warton et al. 2016), et présentent des avantages : la facilité d'interprétation (Warton et al. 2016), la rapidité d'exécution, et la possibilité d'aisément utiliser d'autres outils tels que le partitionnement de la variance (Peres-Neto et al. 2006).

Parmi les avancées technologiques les plus prometteuses pour l'écologie du paysage, figurent celles qui utilisent des algorithmes d'approximation numérique probabilistes et/ou des algorithmes d'intelligence artificielle. Nombreuses sont celles qui utilisent la méthode de Monte-Carlo par chaînes de Markov (MCMC) pour réaliser leurs estimations (Ellison 2004; Michelot et al. 2020; Ponisio et al. 2020), par exemple les modèles gaussiens latents (Fahrmeir & Lang 2001; Rasmussen & Williams 2006; Rue et al. 2009). L'approximation rapide des MCMC par la méthode de Laplace (Beguin et al. 2012; Bolin & Lindgren 2011; Lindgren et al. 2011; Martins et al. 2013; Rue et al. 2017) a, par exemple, été utilisée pour la modélisation de la distribution de plusieurs espèces (Albery et al. 2019), des prises accessoires (Breivik et al. 2017), des phénomènes extrêmes (Opitz 2017; Opitz et al. 2018), de l'erreur liée à l'échantillonnage (Nicolau et al. 2020), des effets génétiques (Mathew et al. 2016; Nustad et al. 2018; Rousset & Ferdy 2014; Selle et al. 2019), et de la chronologie des effets du paysage sur les populations (Leblois et al. 2014). J'ai utilisé un algorithme évolutionnaire dans le Chapitre 3, me permettant ainsi de pallier l'utilisation de valeurs de résistance prédéterminées sur la base d'opinion d'experts, une approche peu performante, mais il existe d'autres algorithmes prometteurs (Peterman & Pope 2021). L'apprentissage par renforcement de gradient (Chen & Guestrin 2016) est très utile pour estimer les propriétés du paysage (Han et al. 2019; Muñoz-Mas et al. 2019). Utiliser une combinaison de modèles plutôt que d'en choisir un seul est avantageux en écologie spatiale (Dormann et al. 2018; Kuczynski et al. 2018). L'apprentissage ensembliste est basé sur cette idée et plusieurs développements récents comme les arbres de régressions additifs bayésiens (Hill et al. 2020; Sparapani et al. 2021) ajoutent beaucoup de flexibilité. Ils ont été utilisés pour mieux prédire la distribution d'espèces végétales (Konowalik & Nosol 2021; Plant et al. 2021), les hôtes de coronavirus (Becker et al. 2020), les refuges face à une autre pandémie (Tytar et al. 2021), ou pour améliorer les études d'association pangénomique (Sawitri et al. 2020), utiles pour la conservation.

Une urgente priorité de recherche en écologie du paysage est d'améliorer la transférabilité des modèles spatiaux entre différents paysages, notamment dans le contexte de changement global. Ceci fut une des plus importantes limites rencontrées dans les Chapitres 2 et 3. Les processus écologiques existent à des échelles spatio-temporelles multiples (Levin 1992; Turner 1989; Wiens 1989), et l'intégration multi-échelles est donc souhaitable pour obtenir de meilleurs modèles (DeCesare et al. 2012; Hobbs 2003; Turner et al. 1989), notamment grâce à une meilleure inclusion de l'incertitude (Barry & Elith 2006; Elith & Leathwick 2009; Johnson & Gillingham 2008). Cependant, les patrons étudiés dans un paysage sont formés dans des contextes différents dans d'autres paysages, donc les conclusions des études scientifiques sont très souvent limitées à une échelle et un contexte écologique à un moment donné. Par exemple, les changements de grain et d'étendue (Anderson et al. 2010; Dale & Fortin 2014) influencent les mesures structurelles du paysage (Wu 2004) et les résultats des modèles basés sur les données de présence-absence (Bean et al. 2014; Guisan et al. 2007; Martin & Fahrig 2012), d'abondance (Fisher et al. 2011; Toews et al. 2017a) ou de flux de gènes (Cushman & Landguth 2010; Galpern et al. 2012). Les différences écologiques mesurées entre paysages (Steinbauer et al. 2012), taxons (Hulme 2008), ou morphes (Kirchheimer et al. 2016), sont parfois attribuables à des artefacts statistiques. L'invariance des relations écologiques, sur des étendues et des grains différents, est nécessaire pour adéquatement transférer des modèles à des fins de conservation, qu'il s'agisse de transférer des modèles obtenus dans des zones d'études limitées/avec des grains de petite taille à des zones plus grandes (« scaling up »; Denny & Benedetti-Cecchi 2012; Morales & Ellner 2002), ou d'utiliser des modèles régionaux pour modéliser un phénomène local (« scaling down »; Maréchaux et al. 2017, Peterson et al. 2018). Cette propriété n'est généralement pas satisfaite et ces deux approches sont donc souvent limitées (Bean et al. 2014; Manzoor et al. 2018; Wenger & Olden 2012; Yates et al. 2018). Une solution partielle à ce problème est une réplication spatiale stratifiée, qui permettrait d'identifier si un élément du paysage a un effet constant dans plusieurs zones d'études, et à plusieurs échelles, ou dans le cas contraire, mieux comprendre comment il varie selon l'échelle et le contexte écologique.

Une autre direction future pour la recherche en écologie du paysage, est de mieux intégrer les propriétés physiques des processus spatiaux et notamment leur dépendance directionnelle, ou anisotropie. La vaste majorité des modèles de connectivité, tels que ceux utilisés dans le Chapitre 3, ne respectent pas l'anisotropie dans les mouvements du vivant et, par exemple, modélisent la dispersion des individus comme ayant la même probabilité dans toutes les directions (Dale & Fortin 2014). Cependant, les mouvements anisotropes sont partout dans la nature. Ils peuvent provenir des interactions sociales (Holt 1996; Pulliam 1988), des gradients d'hétérogénéité spatiale (Benard & McCauley 2008; Bjornstad et al. 2002; Gammon & Maurer 2002; Morin et al. 2009; Trotter et al. 2019), et des forces directionnelles telles que le vent, la gravité, ou les courants. Nonobstant une attention académique grandissante à propos des mouvements anisotropes (Holderegger & Gugerli 2012; Paz-Vinas et al. 2013, 2015), la mauvaise intégration de l'anisotropie affecte l'interprétation des modèles de génétique du paysage, qui visent à déchiffrer le mouvement d'une espèce dans le paysage (Beger et al. 2010; Blanchet et al. 2010; Gaines et al. 2010; Treml et al. 2008). Peu de connaissances sont disponibles sur la façon de relier les matrices asymétriques de prédicteurs et de réponses. Jusqu'ici, les approches se confinèrent aux comparaisons qualitatives (Andreasen et al. 2012) ou à des méthodes ne permettant pas la prédiction, critiquées, et non testées pour cette utilisation (Balkenhol et al. 2009; Guillot & Rousset 2013; Legendre & Fortin 2010; Legendre et al. 2015; Meirmans 2012; Storfer et al. 2010). J'ai commencé à développer une solution (Wittische & James ; en préparation) basée sur les modèles MLPE (Clarke et al. 2002; Peterman 2018; Row et al. 2017; Van Strien et al. 2012), qui prend en compte la source et la destination des migrants de chacune des paires de populations, et permet de distinguer les effets isotropes et anisotropes. Ceci permettrait à terme de mieux prédire la connectivité future, en utilisant, par exemple, les résultats de modèles biophysiques et les variables environnementales classiques comme prédicteurs dans un même modèle de connectivité fonctionnelle (Landguth et al. 2017c; Segura-García et al. 2019). L'évaluation de cette approche à l'aide de simulations réalistes constitue la prochaine étape de ce projet, qui contribuera à une modélisation plus réaliste de ces écosystèmes complexes.

L'écologie du paysage est une science de la complexité. Que ce soit à travers une meilleure prise en compte de l'espace et du temps, par la préservation des propriétés de la communauté et du paysage dans les analyses, ou grâce à l'utilisation de l'intelligence artificielle, nous pouvons capturer cette complexité. Cela est plus que jamais nécessaire pour répondre aux nombreuses questions fondamentales et appliquées, associées avec un changement global encore plus intense que prévu.

RÉFÉRENCES BIBLIOGRAPHIQUES

Abib TH, Chasmer L, Hopkinson C, Mahoney C, Rodriguez LCE. 2019. Seismic line impacts on proximal boreal forest and wetland environments in Alberta. *Sci. Total Environ.* 658:1601–13

ABMI. 2010. *Wall-to-Wall Human Footprint Inventory*. Alberta Biodiversity Monitoring Institute. https://abmi.ca/home/data-analytics/da-top/da-product-overview/Human-Footprint-Products

Abrahms B, Sawyer SC, Jordan NR, McNutt JW, Wilson AM, Brashares JS. 2017. Does wildlife resource selection accurately inform corridor conservation? *J. Appl. Ecol.* 54(2):412–22

Acot P, Blandin P, Hamm BP, Drouin J-M, Blondel-Mégrelis M, et al. 1998. *The European Origins of Scientific Ecology (1800-1901)*. London, UK: Routledge

Adams CIM, Knapp M, Gemmell NJ, Jeunen GJ, Bunce M, et al. 2019. Beyond biodiversity: Can environmental DNA (eDNA) cut it as a population genetics tool? *Genes.* 10(3):192

Aeschbacher S, Selby JP, Willis JH, Coop GM. 2017. Population-genomic inference of the strength and timing of selection against gene flow. *Proc. Natl. Acad. Sci. U. S. A.* 114(27):7061–66

Agapow PM, Burt A. 2001. Indices of multilocus linkage disequilibrium. *Mol. Ecol. Notes*. 1(1–2):101–2

Aguillon SM, Fitzpatrick JW, Bowman R, Schoech SJ, Clark AG, et al. 2017. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genet.* 13(8):1–27

Ainslie B, Jackson PL. 2011. Investigation into mountain pine beetle above-canopy dispersion using weather radar and an atmospheric dispersion model. *Aerobiologia.* 27(1):51–65

Albery GF, Becker DJ, Kenyon F, Nussey DH, Pemberton JM. 2019. The Fine-Scale Landscape of Immunity and Parasitism in a Wild Ungulate Population. *Integr. Comp. Biol.* 59(5):1165–75

Albrechtsen A, Nielsen FC, Nielsen R. 2010. Ascertainment biases in SNP chips affect measures of population divergence. *Mol. Biol. Evol.* 27(11):2534–47

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11(10):697–709

Allio R, Donega S, Galtier N, Nabholz B. 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. *Mol. Biol. Evol.* 34(11):2762–72

Almany GR, Planes S, Thorrold SR, Berumen ML, Bode M, et al. 2017. Larval fish dispersal in a coral-reef seascape. *Nat. Ecol. Evol.* 1(6):1–7

Amman GD. 1972. Mountain Pine Beetlel Brood Production in Relation to Thickness of Lodgepole Pine Phloem. *J. Econ. Entomol.* 65(1):138–40

Amman GD, Baker BH, Stipe LE. 1973. Lodgepole pine losses to mountain pine beetle related to elevation. Victoria, British Columbia

Anderson CD, Epperson BK, Fortin M-J, Holderegger R, James PMA, et al. 2010. Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol. Ecol.* 19(17):3565–75

Andreasen AM, Stewart KM, Longland WS, Beckmann JP, Forister ML. 2012. Identification of source-sink dynamics in mountain lions of the Great Basin. *Mol. Ecol.* 21(23):5689–5701

Apodaca JJ, Trexler JC, Jue NK, Schrader M, Travis J. 2013. Large-scale natural disturbance alters genetic population structure of the sailfin molly, poecilia latipinna. *Am. Nat.* 181(2):254–63

Aukema BH, Carroll AL, Zhu J, Raffa KF, Sickley TA, Taylor SW. 2006. Landscape level analysis of mountain pine beetle in British Columbia, Canada: Spatiotemporal development and spatial synchrony within the present outbreak. *Ecography (Cop.).* 29(3):427–41

Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B. 2000. Effects of colonization processes on genetic diversity: Differences between annual plants and tree species. *Genetics*. 154(3):1309–21

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C. 2013. Individual dispersal, landscape connectivity and ecological networks. *Biol. Rev.* 88(2):310–26

Bailleul D, Mackenzie A, Sacchi O, Poisson F, Bierne N, Arnaud-Haond S. 2018. Large-scale genetic panmixia in the blue shark (Prionace glauca): A single worldwide population, or a genetic lag-time effect of the “grey zone” of differentiation? *Evol. Appl.* 11(5):614–30

Bakker VJ, Van Vuren DH. 2004. Gap-crossing decisions by the red squirrel, a forest-dependent small mammal. *Conserv. Biol.* 18(3):689–97

Balée W, Erickson CL. 2016. *Time, Complexity, and Historical Ecology*. New York, New York, USA: Columbia University Press

Balkenhol N, Cushman SA, Storfer AT, Waits L. 2015. *Landscape Genetics: Concepts, Methods, Applications*. Wiley-Blackwell

Balkenhol N, Waits LP, Dezzani RJ. 2009. Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography (Cop.).* 32(5):818–30

Banks SC, Cary GJ, Smith AL, Davies ID, Driscoll DA, et al. 2013. How does ecological disturbance influence genetic diversity? *Trends Ecol. Evol.* 28(11):670–79

Barbato M, Orozco-terWengel P, Tapio M, Bruford MW. 2015. SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* 6(MAR):1–6

Barker JSF. 2011. Effective population size of natural populations of Drosophila buzzatii, with a comparative evaluation of nine methods of estimation. *Mol. Ecol.* 20(21):4452–71

Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, et al. 2011. Has the Earth's sixth mass extinction already arrived? *Nature*. 471(7336):51–57

Barry S, Elith J. 2006. Error and uncertainty in habitat models. *J. Appl. Ecol.* 43(3):413–23

Bascompte J, Sole R V. 1996. Habitat Fragmentation and Extinction Thresholds in Spatially Explicit Models. *J. Anim. Ecol.* 65(4):465

Baselga A, Araújo MB. 2009. Individualistic vs community modelling of species distributions under climate change. *Ecography.* 32(1):55–65

Batista PD, Janes JK, Boone CK, Murray BW, Sperling FAH. 2016. Adaptive and neutral markers both show continent-wide population structure of mountain pine beetle (Dendroctonus ponderosae). *Ecol. Evol.* 6(17):6292–6300

Battey CJ, Ralph PL, Kern AD. 2020. Space is the place: Effects of continuous spatial structure on analysis of population genetic data. *Genetics*. 215(1):193–214

Bauduin S, Cumming SG, St-Laurent M-H, McIntire EJB. 2020. Integrating functional connectivity in designing networks of protected areas under climate change: A caribou case-study. *PLoS One*. 15(9 September):1–19

Bauer S, Shamoun-Baranes J, Nilsson C, Farnsworth A, Kelly JF, et al. 2019. The grand challenges of migration ecology that radar aeroecology can help answer. *Ecography (Cop.).* 42(5):861–75

Bean WT, Prugh LR, Stafford R, Butterfield HS, Westphal M, Brashares JS. 2014. Species distribution models of an endangered rodent offer conflicting measures of habitat quality at multiple scales. *J. Appl. Ecol.* 51(4):1116–25

Beauchesne D, Jaeger JAG, St-Laurent M-H. 2013. Disentangling woodland caribou movements in response to clearcuts and roads across temporal scales. *PLoS One*. 8(11):1–11

Beauchesne D, Jaeger JAG, St-Laurent M-H. 2014. Thresholds in the capacity of boreal caribou to cope with cumulative disturbances: evidence from space use patterns. *Biol. Conserv.* 172:190–99

Becker DJ, Albery GF, Sjodin AR, Poisot T, Dallas TA, et al. 2020. Predicting wildlife hosts of betacoronaviruses for SARS-CoV-2 sampling prioritization. *bioRxiv*

Beerli P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Mol. Ecol.* 13(4):827–36

Beger M, Linke S, Watts M, Game E, Treml EA, et al. 2010. Incorporating asymmetric connectivity into spatial decision making for conservation. *Conserv. Lett.* 3(5):359–68

Beguin J, Martino S, Rue H, Cumming SG. 2012. Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation. *Methods Ecol. Evol.* 3(5):921–29

Beirne C, Sun C, Tattersall ER, Burgar JM, Fisher JT, Burton AC. 2021. Multispecies modelling reveals potential for habitat restoration to re-establish boreal vertebrate community dynamics. *J. Appl. Ecol.* 58(12):2821–32

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F. 2012. Impacts of climate change on the future of biodiversity. *Ecol. Lett.* 15(4):365–77

Bellard C, Cassey P, Blackburn TM. 2016. Alien species as a driver of recent extinctions. *Biol. Lett.* 12(4):20150623

Bellis ED, Graves HB. 1971. Deer mortality on a Pennsylvania interstate highway. *J. Wildl. Manage.* 35(2):232–37

Benard MF, McCauley SJ. 2008. Integrating across Life-History Stages: Consequences of Natal Habitat Effects on Dispersal. *Am. Nat.* 171(5):553–67

Beninde J, Feldmeier S, Werner M, Peroverde D, Schulte U, et al. 2016. Cityscape genetics: structural vs. functional connectivity of an urban lizard population. *Mol. Ecol.* 25(20):4984–5000

Benítez-López A, Alkemade R, Schipper AM, Ingram DJ, Verweij PA, et al. 2017. The impact of hunting on tropical mammal and bird populations. *Science.* 356:180–83

Benítez-López A, Santini L, Schipper AM, Busana M, Huijbregts MAJ. 2019. Intact but empty forests? Patterns of huntinginduced mammal defaunation in the tropics. *PLoS Biol.* 17(5):1–18

Bennett PI, Stone JK. 2019. Environmental variables associated with Nothophaeocryptopus gaeumannii population structure and Swiss needle cast severity in Western Oregon and Washington. *Ecol. Evol.* 9(19):11379–94

Benson JF, Patterson B. 2013. Moose (Alces alces) predation by eastern coyotes (Canis latrans) and eastern coyote × eastern wolf (Canis latrans × Canis lycaon) hybrids. *Can. J. Zool.* 91:837–41

Bentz BJ, Mullins DE. 1999. Ecology of Mountain Pine Beetle (Coleoptera: Scolytidae) Cold Hardening in the Intermountain West. *Environ. Entomol.* 28(4):577–87

Bentz BJ, Powell JA. 2014. Mountain pine beetle seasonal timing and constraints to bivoltinism: A comment on Mitton and Ferrenberg, "Mountain pine beetle develops an unprecedented summer generation in response to climate warming." *Am. Nat.* 184(6):787–96

Bentz BJ, Régnière J, Fettig CJ, Hansen EM, Hayes JL, et al. 2010. Climate Change and Bark Beetles of the Western United States and Canada: Direct and Indirect Effects. *Bioscience.*

60(8):602–13

Berg EE, David Henry J, Fastie CL, De Volder AD, Matsuoka SM. 2006. Spruce beetle outbreaks on the Kenai Peninsula, Alaska, and Kluane National Park and Reserve, Yukon Territory: Relationship to summer temperatures and regional differences in disturbance regimes. *For. Ecol. Manage.* 227(3 SPEC. ISS.):219–32

Bezault E, Balaresque P, Toguyeni A, Fermon Y, Araki H, et al. 2011. Spatial and temporal variation in population genetic structure of wild Nile tilapia (Oreochromis niloticus) across Africa. *BMC Genet.* 12(102):

Bezemer N, Krauss SL, Roberts DG, Hopper SD. 2019. Conservation of old individual trees and small populations is integral to maintain species' genetic diversity of a historically fragmented woody perennial. *Mol. Ecol.* 28(14):3339–57

Bhatia G, Patterson N, Sankararaman S, Price AL. 2013. Estimating and interpreting F. *Genome Res.* (2):1–9

Biémont C, Vieira C. 2006. Genetics: Junk DNA as an evolutionary force. *Nature*. 443(7111):521–24

Bjornstad ON, Peltonen M, Liebhold AM, Baltensweiler W. 2002. Waves of Larch Budmoth Outbreaks in the European Alps. *Science.* 298(5595):1020–23

Blanchet FG, Legendre P, Borcard D. 2008. Forward selection of explanatory variables. *Ecology*. 89(9):2623–32

Blanchet S, Rey O, Etienne R, Lek S, Loot G. 2010. Species-specific responses to landscape fragmentation: Implications for management strategies. *Evol. Appl.* 3(3):291–304

Bohl CL, Kass JM, Anderson RP. 2019. A new null model approach to quantify performance and significance for ecological niche models of species distributions. *J. Biogeogr.* 46(6):1101–11

Boisjoly D, Ouellet J-P, Courtois R. 2010a. Coyote habitat selection and management implications for the Gaspésie Caribou. *J. Wildl. Manage.* 74(1):3–11

Boisjoly D, Ouellet J-P, Courtois R. 2010b. Coyote Habitat Selection and Management Implications for the Gaspésie Caribou. *J. Wildl. Manage.* 74(1):3–11

Bolin D, Lindgren FK. 2011. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann. Appl. Stat.* 5(1):523–50

Bolnick DI, Nosil P. 2007. Natural selection in populations subject to a migration load. *Evolution.* 61(9):2229–43

Bond ML, Bradley CM, Kiffner C, Morrison TA, Lee DE. 2017. A multi-method approach to delineate and validate migratory corridors. *Landsc. Ecol.* 32(8):1705–21

Bonte D, Van Dyck H, Bullock JM, Coulon A, Delgado M, et al. 2012. Costs of dispersal. *Biol. Rev. Camb. Philos. Soc.* 87(2):290–312

Borcard D, Gillet F, Legendre P. 2018. *Numerical Ecology with R, Second Edition*. Springer International Publishing AG

Borcard D, Legendre P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Modell.* 153(1–2):51–68

Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H. 2004. Dissecting the spatial structure of ecological data at multiple scales. *Ecology*. 85(7):1826–32

Borcard D, Legendre P, Drapeau P. 1992. Partialling out the spatial component of ecological variation. *Ecology*. 73(3):1045–55

Boucher D, Boulanger Y, Aubin I, Bernier PY, Beaudoin A, et al. 2018. Current and projected cumulative impacts of fire, drought, and insects on timber volumes across Canada. *Ecol. Appl.* 28(5):1245–59

Boulanger Y, Fabry F, Kilambi A, Pureswaran DS, Sturtevant BR, Saint-Amant R. 2017. The use of weather surveillance radar and high-resolution three dimensional weather data to monitor a spruce budworm mass exodus flight. *Agric. For. Meteorol.* 234–235:127–35

Boutin S, Boyce MS, Hebblewhite M, Hervieux D, Knopff KH, et al. 2012. Why are caribou declining in the oil sands? *Front. Ecol. Environ.* 10(2):65–67

Bowman J, Adey E, Angoh SYJ, Baici JE, Brown MGC, et al. 2020. Effects of cost surface uncertainty on current density estimates from circuit theory. *PeerJ*. 8:1–18

Bowman J, Ray JC, Magoun AJ, Johnson DS, Dawson FN. 2010. Roads, logging, and the large-mammal community of an eastern Canadian boreal forest. *Can. J. Zool.* 88(5):454–67

Bradburd GS, Ralph PL. 2019. Spatial Population Genetics: It's About Time. *Annu. Rev. Ecol. Evol. Syst.* 50(1):427–49

Brainerd SM, Rolstad J. 2002. Habitat selection by Eurasian pine martens Martes martes in managed forests of southern boreal Scandinavia. *Wildlife Biol.* 8(4):289–97

Breivik ON, Storvik G, Nedreaas K. 2017. Latent Gaussian models to predict historical bycatch in commercial fishery. *Fish. Res.* 185:62–72

Brennan A, Beytell P, Aschenborn O, Du Preez P, Funston PJ, et al. 2020. Characterizing multispecies connectivity across a transfrontier conservation landscape. *J. Appl. Ecol.* 57(9):1700–1710

Broadley K, Burton AC, Avgar T, Boutin S. 2019. Density-dependent space use affects interpretation of camera trap detection rates. *Ecol. Evol.* 9(24):14031–41

Broquet T, Petit EJ. 2009. Molecular Estimation of Dispersal for Ecology and Population Genetics. *Annu. Rev. Ecol. Evol. Syst.* 40(1):193–216

Burgar JM, Burton AC, Fisher JT. 2019. The importance of considering multiple interacting species for conservation of species at risk. *Conserv. Biol.* 33(3):709–15

Burgess D, Freeling M. 2014. The most deeply conserved noncoding sequences in plants serve similar functions to those in vertebrates despite large differences in evolutionary rates. *Plant Cell*. 26(3):946–61

Burton AC, Neilson EW, Moreira D, Ladle A, Steenweg R, et al. 2015. Wildlife camera trapping: A review and recommendations for linking surveys to ecological processes. *J. Appl. Ecol.* 52(3):675–85

Buschbom J, Yanbaev Y, Degen B. 2011. Efficient long-distance gene flow into an isolated relict oak stand. *J. Hered.* 102(4):464–72

Butchart SHM, Walpole M, Collen B, von Strien A, Scharlemann JPW, et al. 2010. Global biodiversity : indicators of recent declines. *Science.* 328(5982):1164–68

Capblancq T, Fitzpatrick MC, Bay RA, Exposito-Alonso M, Keller SR. 2020. Genomic Prediction of (Mal)Adaptation across Current and Future Climatic Landscapes. *Annu. Rev. Ecol. Evol. Syst.* 51(1):245–69

Caravaggi A, Banks PB, Burton AC, Finlay CMV, Haswell PM, et al. 2017. A review of camera trapping for conservation behaviour research. *Remote Sens. Ecol. Conserv.* 3(3):109–22

Carlos-Júnior LA, Creed JC, Marrs R, Lewis RJ, Moulton TP, et al. 2020. Generalized Linear Models outperform commonly used canonical analysis in estimating spatial structure of presence/absence data. *PeerJ*. 8:e9777

Carpi LC, Schieber TA, Pardalos PM, Marfany G, Masoller C, et al. 2019. Assessing diversity in multiplex networks. *Sci. Rep.* 9(1):1–12

Carroll AL, Taylor SW, Régnière J, Safranyik L. 2004. Effects of Climate Change on Range Expansion by the Mountain Pine Beetle in British Columbia. *Mt. Pine Beetle Symp. Challenges Solut. Oct. 30-31, 2003, Kelowna, Br. Columbia*, pp. 223–32. Victoria, British Columbia, Canada: Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada

Catano CP, Dickson TL, Myers JA. 2017. Dispersal and neutral sampling mediate contingent effects of disturbance on plant beta-diversity: a meta-analysis. *Ecol. Lett.* 20(3):347–56

Cavada N, Havmøller RW, Scharff N, Rovero F. 2019. A landscape-scale assessment of tropical mammals reveals the effects of habitat and anthropogenic disturbance on community occupancy. *PLoS One*. 14(4):1–15

Caye K, Deist TM, Martins H, Michel O, François O. 2016. TESS3: Fast inference of spatial population structure and genome scans for selection. *Mol. Ecol. Resour.* 16(2):540–48

Cayuela H, Rougemont Q, Prunier JG, Moore JS, Clobert J, et al. 2018. Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Mol. Ecol.* 27(20):3976–4010

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. 2015. Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Sci. Adv.* 1(5):e1400253

Cederlund GN, Okarma H. 1988. Home range and habitat use of adult female moose. *J. Wildl. Manage.* 52(2):336–43

Chandler RB, Royle AJ. 2013. Spatially explicit models for inference about density in unmarked or partially marked populations. *Ann. Appl. Stat.* 7(2):936–54

Charlesworth D, Willis JH. 2009. The genetics of inbreeding depression. *Nat. Rev. Genet.* 10(11):783–96

Charney ND. 2012. Evaluating expert opinion and spatial scale in an amphibian model. *Ecol. Modell.* 242:37–45

Chen T, Guestrin C. 2016. XGBoost: A scalable tree boosting system. *Proc. ACM SIGKDD Int.*

Conf. Knowl. Discov. Data Min., pp. 785–94

Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R. 2005. Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res.* 15(11):1496–1502

Clark JS, Nemergut D, Seyednasrollah B, Turner PJ, Zhang S. 2017. Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecol. Monogr.* 87(1):34–56

Clarke RT, Rothery P, Raybould AF. 2002. Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *J. Agric. Biol. Environ. Stat.* 7(3):361–72

Clobert J, Le Galliard J-F, Cote J, Meylan S, Massot M. 2009. Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecol. Lett.* 12(3):197–209

Condit R, Pitman N, Leigh EG, Chave J, Terborgh J, et al. 2002. Beta-diversity in tropical forest trees. *Science.* 295(5555):666–69

Conradi T, Temperton VM, Kollmann J. 2017. Beta diversity of plant species in human-transformed landscapes: Control of community assembly by regional productivity and historical connectivity. *Perspect. Plant Ecol. Evol. Syst.* 24:1–10

Cooke BJ, Carroll AL. 2017. Predicting the risk of mountain pine beetle spread to eastern pine forests: Considering uncertainty in uncertain times. *For. Ecol. Manage.* 396:11–25

Copas JB, Long T. 1991. Estimating the Residual Variance in Orthogonal Regression with

Variable Selection. *Stat.* 40(1):51

Correa Ayram CA, Mendoza ME, Etter A, Salicrup DRP. 2016. Habitat connectivity in biodiversity conservation: A review of recent studies and applications. *Prog. Phys. Geogr.* 40(1):7–37

Côté IM, Darling ES, Brown CJ. 2016. Interactions among ecosystem stressors and their importance in conservation. *Proc. R. Soc. B Biol. Sci.* 283(1824):1–9

Côté SD, Rooney TP, Tremblay JP, Dussault C, Waller DM. 2004. Ecological impacts of deer overabundance. *Annu. Rev. Ecol. Evol. Syst.* 35:113–47

Coulon A, Morellet N, Goulard M, Cargnelutti B, Angibault JM, Hewison AJM. 2008. Inferring the effects of landscape structure on roe deer (Capreolus capreolus) movements using a step selection function. *Landsc. Ecol.* 23(5):603–14

Courbin N, Fortin D, Dussault C, Courtois R. 2009. Landscape management for woodland caribou: The protection of forest blocks influences wolf-caribou co-occurrence. *Landsc. Ecol.* 24(10):1375–88

Courbin N, Fortin D, Dussault C, Courtois R. 2014. Logging-induced changes in habitat network connectivity shape behavioral interactions in the wolf – caribou – moose system. *Ecol. Monogr.* 84(2):265–85

Courtois R, Ouellet J-P, Breton L, Gingras A, Dussault C. 2007. Effects of forest disturbance on density, space use, and mortality of woodland caribou. *Ecoscience.* 14(4):491–98

Crête M, Desrosiers A. 1995. Range expansion of coyotes, Canis latrans, threatens a remnant herd of caribou, Rangifer tarandus, in southeastern Québec. *Can. Field-Naturalist*. 109:227–235

Crête M, Drolet B, Huot J, Fortin M, Doucet G. 1995. Postfire sequence of emerging diversity among mammals and birds in the north of the boreal forest in Quebec. *Can. J. For. Res.* 25:1509–18

Cubry P, Vigouroux Y, François O. 2017. The Empirical Distribution of Singletons for Geographic Samples of DNA Sequences. *Front. Genet.* 8:139

Culina A, van den Berg I, Evans S, Sánchez-Tójar A. 2020. Low availability of code in ecology: A call for urgent action. *PLoS Biol.* 18(7):1–9

Cullingham CI, Cooke JEK, Dang S, Davis CS, Cooke BJ, Coltman DW. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. *Mol. Ecol.* 20(10):2157–71

Cushman SA, Landguth EL. 2010. Scale dependent inference in landscape genetics. *Landsc. Ecol.* 25(6):967–79

Dale MRT, Fortin M-J. 2010. From Graphs to Spatial Graphs. *Annu. Rev. Ecol. Evol. Syst.* 41(1):21–38

Dale MRT, Fortin M-J. 2014. *Spatial Analysis: A Guide for Ecologists*. Cambridge, UK: Cambridge University Press. 2nd ed.

Darimont CT, Fox CH, Bryan HM, Reimchen TE. 2015. The unique ecology of human predators. *Science.* 349(6250):858–60

Darlington S. 2018. *Cumulative Effects of Human Landscape Change, Predators, and Natural Habitat Drive Distrubitions of an Invasive Ungulate*. University of Victoria

Darwin CR, Wallace AR. 1858. On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *Zool. J. Linn. Soc.* 3(9):45–62

De Cáceres M, Coll L, Legendre P, Allen RB, Wiser SK, et al. 2019. Trajectory analysis in community ecology. *Ecol. Monogr.* 89(2):1–20

de la Giroday H-MC, Carroll AL, Aukema BH. 2012. Breach of the northern Rocky Mountain geoclimatic barrier: Initiation of range expansion by the mountain pine beetle. *J. Biogeogr.* 39(6):1112–23

de la Giroday H-MC, Carroll AL, Lindgren BS, Aukema BH. 2011. Incoming! Association of landscape features with dispersing mountain pine beetle populations during a range expansion event in western Canada. *Landsc. Ecol.* 26(8):1097–1110

DeCesare NJ, Hebblewhite M, Schmiegelow F, Hervieux D, McDermid GJ, et al. 2012. Transcending scale dependence in identifying habitat with resource selection functions. *Ecol. Appl.* 22(4):1068–83

DeMars CA, Boutin S. 2018. Nowhere to hide: Effects of linear features on predator–prey dynamics in a large mammal system. *J. Anim. Ecol.* 87(1):274–84

Dénes F V., Silveira LF, Beissinger SR. 2015. Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. *Methods Ecol. Evol.* 6(5):543–56

Denny M, Benedetti-Cecchi L. 2012. Scaling up in ecology: Mechanistic approaches. *Annu. Rev. Ecol. Evol. Syst.* 43:1–22

Derksen S, Keselman HJ. 1992. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *Br. J. Math. Stat. Psychol.* 45(2):265–82

Devictor V, Julliard R, Jiguet F, Devictor V, Julliard R, Jiguet F. 2008. Distribution of Specialist and Generalist Species along Spatial Gradients of Habitat Disturbance and Fragmentation. *Oikos*. 117(4):507–14

Dickie M, Serrouya R, McNay RS, Boutin S. 2016. Faster and farther: wolf movement on linear features and implications for hunting behaviour. *J. Appl. Ecol.* 54(1):253–63

Dickson BG, Albano CM, Anantharaman R, Beier P, Fargione J, et al. 2019. Circuit-theory applications to connectivity science and conservation. *Conserv. Biol.* 33(2):239–49

Diehr G, Hoflin DR. 1974. Approximating the distribution of the sample R2 in best subset regressions. *Technometrics*. 16(2):317–20

Díez-del-Molino D, Sánchez-Barreiro F, Barnes I, Gilbert MTP, Dalén L. 2018. Quantifying Temporal Genomic Erosion in Endangered Species. *Trends Ecol. Evol.* 33(3):176–85

Dileo MF, Siu JC, Rhodes MK, Lõpez-Villalobos A, Redwine A, et al. 2014. The gravity of pollination: Integrating at-site features into spatial analysis of contemporary pollen movement. *Mol. Ecol.* 23(16):3973–82

DiLeo MF, Wagner HH. 2016. A Landscape Ecologist's Agenda for Landscape Genetics. *Curr.*

Landsc. Ecol. Reports. 1(3):115–26

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B. 2014. Defaunation in the Anthropocene. *Science.* 401(6195):401–6

Dise NB, Ashmore M, Belyazid S, Bleeker A, Bobbink R, et al. 2011. Nitrogen as a threat to European soil quality. In *The European Nitrogen Assessment*, eds. MA Sutton, CM Howard, JW Erisman, G Billen, A Bleeker, et al., pp. 495–510. Cambridge, UK: Cambridge University Press

Doherty TS, Dickman CR, Nimmo DG, Ritchie EG. 2015. Multiple threats, or multiplying the threats? Interactions between invasive predators and other ecological disturbances. *Biol. Conserv.* 190:60–68

Dormann CF, Calabrese JM, Guillera-Arroita G, Matechou E, Bahn V, et al. 2018. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecol. Monogr.* 88(4):485–504

Dornelas M, Magurran AE, Buckland ST, Chao A, Chazdon RL, et al. 2013. Quantifying temporal change in biodiversity: Challenges and opportunities. *Proc. R. Soc. B Biol. Sci.* 280(1750):1–10

Doyle PG, Snell JL. 1984. *Random Walks and Electric Networks*. Washington DC, USA: Mathematical Association of America

Draheim HM, Moore JA, Fortin M-J, Scribner KT. 2018. Beyond the snapshot: Landscape genetic analysis of time series data reveal responses of American black bears to landscape change. *Evol. Appl.* 11(8):1219–30

Dray S, Bauman D, Blanchet FG, Borcard D, Clappe S, et al. 2019. adespatial: multivariate multiscale spatial analysis. R package version 0.3-7. https://cran.r-project.org/package=adespatial

Dray S, Legendre P, Peres-Neto PR. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Modell.* 196(3–4):483–93

Dufresnes C, Pellet J, Bettinelli-Riccardi S, Thiébaud J, Perrin N, Fumagalli L. 2016. Massive genetic introgression in threatened northern crested newts (Triturus cristatus) by an invasive congener (T. carnifex) in Western Switzerland. *Conserv. Genet.* 17(4):839–46

Dulvy NK, Yvonne S, Reynolds JD. 2003. Extinction vulnerability in marine populations. *Fish Fish.* 4(1):25–64

Duning JB. 1995. Spatially explicit population models: current forms and future uses. *Ecol. Appl.* 5(1):3–11

Ebach MC, Goujet DF. 2006. The first biogeographical map. *J. Biogeogr.* 33(5):761–69

Egerton FN. 2013. History of Ecological Sciences, Part 47: Ernst Haeckel's Ecology. *Bull. Ecol. Soc. Am.* 94(3):222–44

Ehrlich PR, Raven PH. 1964. Butterflies and Plants : A Study in Coevolution. *Evolution.* 18(4):586–608

Eldon B, Birkner M, Blath J, Freund F. 2015. Can the site-frequency spectrum distinguish exponential population growth from multiple-merger coalescents? *Genetics*. 199(3):841–56

Elith J, Leathwick JR. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* 40:677–97

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. 2011. A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17(1):43–57

Ellison AM. 2004. Bayesian inference in ecology. *Ecol. Lett.* 7(6):509–20

Ellstrand NC, Elam DR. 1993. Population genetic conseqences of small population size: Implications for plant conservation. *Annu. Rev. Ecol. Syst.* 24:217–42

Elton CS. 1927. *Animal Ecology*. London, UK: William Clowes & Sons

Epperson BK, McRae BH, Scribner K, Cushman SA, Rosenberg MS, et al. 2010. Utility of computer simulations in landscape genetics. *Mol. Ecol.* 19(17):3549–64

Epps CW, Wasser SK, Keim JL, Mutayoba BM, Brashares JS. 2013. Quantifying past and present connectivity illuminates a rapidly changing landscape for the African elephant. *Mol. Ecol.* 22(6):1574–88

Estes L, Elsen PR, Treuer T, Ahmed L, Caylor K, et al. 2018. The spatial and temporal domains of modern ecology. *Nat. Ecol. Evol.* 2(5):819–26

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013. Robust Demographic Inference from Genomic and SNP Data. *PLoS Genet.* 9(10):e1003905

Fahrig L. 2003a. Effects of habitat fragmentation on biodiversity. *Annu. Rev. Ecol. Evol. Syst.*

34(1):487–515

Fahrig L. 2003b. Effects of Habitat Fragmentation on Biodiversity. *Annu. Rev. Ecol. Evol. Syst.* 34(1):487–515

Fahrmeir L, Lang S. 2001. Bayesian inference for generalized additive mixed models based on Markov random field priors. *J. R. Stat. Soc. Ser. C Appl. Stat.* 50(2):201–20

Fall A, Fortin M-J, Manseau M, O'Brien D. 2007. Spatial graphs: Principles and applications for habitat connectivity. *Ecosystems*. 10(3):448–61

Fenderson LE, Kovach AI, Llamas B. 2019. Spatiotemporal landscape genetics: Investigating ecology and evolution through space and time. *Mol. Ecol.* 29(2):218–46

Fenderson LE, Kovach AI, Llamas B. 2020. Spatiotemporal landscape genetics: Investigating ecology and evolution through space and time. *Mol. Ecol.* 29(2):218–46

Festa-Bianchet M, Ray JC, Boutin S, Côté SD, Gunn A. 2011. Conservation of caribou (rangifer tarandus) in Canada: An uncertain future 1. *Can. J. Zool.* 89(5):419–34

Fisher JT, Anholt B, Volpe JP. 2011. Body mass explains characteristic scales of habitat selection in terrestrial mammals. *Ecol. Evol.* 1(4):517–28

Fisher JT, Burton AC. 2018. Wildlife winners and losers in an oil sands landscape. *Front. Ecol. Environ.* 16(6):323–28

Fisher JT, Burton AC, Nolan L, Hiltz M, Roy LD. 2016. White-tailed deer distribution, density, and habitat selection in the northeast boreal forest. Alberta Innovates -Technology Futures, Vegreville, Alberta.

Fisher JT, Burton AC, Nolan L, Roy L. 2020. Influences of landscape change and winter severity on invasive ungulate persistence in the Nearctic boreal forest. *Sci. Rep.* 10(1):1–11

Fisher JT, Wilkinson L. 2005. The response of mammals to forest fire and timber harvest in the North American boreal forest. *Mamm. Rev.* 35(1):51–81

Fisher KE, Dixon PM, Han G, Adelman JS, Bradbury SP. 2021. Locating large insects using automated VHF radio telemetry with a multi-antennae array. *Methods Ecol. Evol.* 12(3):494–506

Fisher MC, Garner TWJ. 2020. Chytrid fungi and global amphibian declines. *Nat. Rev. Microbiol.* 18(6):332–43

Foley JA, DeFries R, Asner GP, Barford C, Bonan G, et al. 2005. Global consequences of land use. *Science.* 309(5734):570–74

Forbes GJ, Theberge JB. 1993. Multiple Landscape Scales and Winter Distribution of Moose, Alces alces, in a Forest Ecotone. *Can. Field-Naturalist*. 107(2):201–7

Frantz AC, Pourtois JT, Heuertz M, Schley L, Flamand MC, et al. 2006. Genetic structure and assignment tests demonstrate illegal translocation of red deer (Cervus elaphus) into a continuous population. *Mol. Ecol.* 15(11):3191–3203

Frantz AC, Zachos FE, Bertouille S, Eloy MC, Colyn M, Flamand MC. 2017. Using genetic tools to estimate the prevalence of non-native red deer (Cervus elaphus) in a Western

European population. *Ecol. Evol.* 7(19):7650–60

Freedman LS, Pee D, Midthune DN. 1992. The Problem of Underestimating the Residual Error Variance in Forward Stepwise Regression. *Stat.* 41(4):405

Frenette J, Pelletier F, St-Laurent M-H. 2020. Linking habitat, predators and alternative prey to explain recruitment variations of an endangered caribou population. *Glob. Ecol. Conserv.* 22:e00920

Gaines SD, White C, Carr MH, Palumbi SR. 2010. Designing marine reserve networks for both conservation and fisheries management. *Proc. Natl. Acad. Sci.* 107(43):18286–93

Gallagher R, Appenzeller T. 1999. Beyond Reductionism. *Science.* 284:79

Galpern P, Manseau M, Wilson P. 2012. Grains of connectivity: Analysis at multiple spatial scales in landscape genetics. *Mol. Ecol.* 21(16):3996–4009

Gammon DE, Maurer BA. 2002. Evidence for non-uniform dispersal in the biological invasions of two naturalized North American bird species. *Glob. Ecol. Biogeogr.* 11(2):155–61

Gardner RH, Urban DL. 2007. Neutral models for testing landscape hypotheses. *Landsc. Ecol.* 22(1):15–29

Gattepaille LM, Jakobsson M, Blum MGB. 2013. Inferring population size changes with sequence and SNP data: Lessons from human bottlenecks. *Heredity (Edinb).* 110(5):409–19

Gause GF. 1934. *The Struggle for Existence*. Baltimore, USA: Williams and Wilkins Company

Gavin MC, Solomon JN, Blank SG. 2010. Measuring and monitoring illegal use of natural resources. *Conserv. Biol.* 24(1):89–100

Gibson L, Lynam AJ, Bradshaw CJA, He F, Bickford DP, et al. 2013. Near-complete extinction of native small mammal fauna 25 years after forest fragmentation. *Science.* 341(6153):1508–10

Göbel F, Jagers AA. 1974. Random walks on graphs. *Stoch. Process. their Appl.* 2(4):311–36

Goodall DW. 1954. Objective methods for the classification of vegetation. *Aust. J. Bot.* 2(3):304–24

Gotelli NJ. 2001. Research frontiers in null model analysis. *Glob. Ecol. Biogeogr.* 10(4):337–43

Gotelli NJ, Ulrich W. 2012. Statistical challenges in null model analysis. *Oikos*. 121(2):171–80

Government of Alberta. 2019. *GeoDiscover Alberta - Alberta's Geospatial Data Authority*. https://geodiscover.alberta.ca/geoportal/rest/find/document?rid=local&ridName=GeoDisco ver Alberta&rids=local&start=1&max=10&orderBy=relevance&dataCategory=transportation&maxSearchTimeMilliSec=10000&g=searchpageresults&f=searchPage

Government of Canada. Species at Risk Act - Statutes of Canada 2002 c. 29 https://laws.justice.gc.ca/eng/acts/S-15.3/

Government of Saskatchewan. 2021. Saskatchewan and Alberta Renew Agreement to Battle Mountain Pine Beetle Threat. Dec. 6, . https://www.saskatchewan.ca/government/news-and-media/2021/december/06/saskatchewan-and-alberta-renew-agreement-to-battle-mountain-pine-beetle-threat

Graham LJ, Haines-Young RH, Field R. 2018. The incidence function model as a tool for landscape-scale ecological impact assessments. *Landsc. Urban Plan.* 170:187–94

Gravel D, Baiser B, Dunne JA, Kopelke JP, Martinez ND, et al. 2019. Bringing Elton and Grinnell together: a quantitative framework to represent the biogeography of ecological interaction networks. *Ecography (Cop.).* 42(3):401–15

Griffith DA, Peres-Neto PR. 2006. Spatial modeling in ecology: The flexibility of eigenfunction spatial analyses. *Ecology*. 87(10):2603–13

Grinnell J. 1917. The Niche-Relationships of the California Thrasher. *Auk*. 34(4):427–33

Guillot G, Rousset F. 2013. Dismantling the Mantel tests. *Methods Ecol. Evol.* 4(4):336–44

Guisan A, Graham CH, Elith J, Huettmann F, Dudik M, et al. 2007. Sensitivity of predictive species distribution models to change in grain size. *Divers. Distrib.* 13(3):332–40

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* 5(10):e1000695

Haddad NM, Brudvig L a., Clobert J, Davies KF, Gonzalez A, et al. 2015. Habitat fragmentation and its lasting impact on Earth’s ecosystems. *Sci. Adv.* 1(2):1–9

Haeckel EHPA. 1866. *Generelle Morphologie Der Organismen. Allgemeine Grundzüge Der*

Organischen Formen-Wissenschaft, Mechanische Begründet Durch Die von Charles Darwin Reformirte Descendenz-Theorie. Volume II: Alllgemeine Entwickelungsgeschichte Der Organismen. Berlin, Germany: Georg Reimer

Haller BC, Messer PW. 2019. SLiM 3: Forward Genetic Simulations Beyond the Wright-Fisher Model. *Mol. Biol. Evol.* 36(3):632–37

Halsey LG, Curran-Everett D, Vowler SL, Drummond GB. 2015. The fickle P value generates irreproducible results. *Nat. Methods*. 12(3):179–85

Halu A, De Domenico M, Arenas A, Sharma A. 2019. The multiplex network of human diseases. *npj Syst. Biol. Appl.* 5:15

Han Y, Wu J, Zhai B, Pan Y, Huang G, et al. 2019. Coupling a Bat Algorithm with XGBoost to Estimate Reference Evapotranspiration in the Arid and Semiarid Regions of China. *Adv. Meteorol.* 9575782

Hanski I. 1994. A practical model of metapopulation dynamics. *J. Anim. Ecol.* 63(1):151–62

Hardy J, Singleton A. 2009. Genomewide Association Studies and Human Disease. *N. Engl. J. Med.* 360(17):1759–68

Harris DJ. 2015. Generating realistic assemblages with a joint species distribution model. *Methods Ecol. Evol.* 6(4):465–73

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P. 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol. Appl.* 7(9):1008–25

Harwood J, Stokes K. 2003. Coping with uncertainty in ecological advice: Lessons from fisheries. *Trends Ecol. Evol.* 18(12):617–22

Hastie T, Tibshirani R, Friedman J. 2017. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer International Publishing AG

Hastings A, Harrison S. 1994. Metapopulation dynamics and genetics. *Annu. Rev. Ecol. Syst.* 25:167–88

Hebblewhite M. 2017. Billion dollar boreal woodland caribou and the biodiversity impacts of the global oil and gas industry. *Biol. Conserv.* 206:102–11

Hedrick PW, Kalinowski ST. 2000. Inbreeding depression in conservation biology. *Annu. Rev. Ecol. Syst.* 31:139–62

Hegyi G, Garamszegi LZ. 2011. Using information theory as a substitute for stepwise regression in ecology and behavior. *Behav. Ecol. Sociobiol.* 65(1):69–76

Heim N, Fisher JT, Clevenger A, Paczkowski J, Volpe J. 2017. Cumulative effects of climate and landscape change drive spatial distribution of Rocky Mountain wolverine (Gulo gulo L.). *Ecol. Evol.* 7(21):8903–14

Heim N, Fisher JT, Volpe J, Clevenger AP, Paczkowski J. 2019. Carnivore community response to anthropogenic landscape change: species-specificity foils generalizations. *Landsc. Ecol.* 34(11):2493–2507

Herfindal I, Tremblay J-P, Hansen BB, Solberg EJ, Heim M, Saether B-E. 2009. Scale dependency and functional response in moose habitat selection. *Ecography.* 32(5):849–59

Hervieux D, Hebblewhite M, Stepnisky D, Bacon M, Boutin S. 2014. Managing wolves (Canis lupus) to recover threatened woodland caribou (Rangifer tarandus caribou) in Alberta. *Can. J. Zool.* 92:1029–37

Hilborn R. 2012. *Overfishing: What Everyone Needs to Know*. New York, USA: Oxford University Press

Hill J, Linero A, Murray J. 2020. Bayesian additive regression trees: A review and look forward. *Annu. Rev. Stat. Its Appl.* 7:251–78

Hirschhorn JN, Daly MJ. 2005. Genome-wide association studies for common diseases and complex traits. *Nat. Rev. Genet.* 6(2):95–108

Hobbie JE, Carpenter SR, Grimm NB, Gosz JR, Seastedt TR. 2003. The US Long Term Ecological Research program. *Bioscience*. 53(1):21–32

Hobbs NT. 2003. Challenges and opportunities in integrating ecological knowledge across scales. *For. Ecol. Manage.* 181(1–2):223–38

Hody JW, Kays R. 2018. Mapping the expansion of coyotes (Canis latrans) across North and central America. *Zookeys*. 2018(759):81–97

Hofmeester TR, Rowcliffe JM, Jansen PA. 2017. A simple method for estimating the effective detection distance of camera traps. *Remote Sens. Ecol. Conserv.* 3(2):81–89

Hogg EH. 1997. Temporal scaling of moisture and the forest-grassland boundary in western Canada. *Agric. For. Meteorol.* 84(1–2):115–22

Holderegger R, Gugerli F. 2012. Where do you come from, where do you go? Directional migration rates in landscape genetics. *Mol. Ecol.* 21(23):5640–42

Holderegger R, Kamm U, Gugerli F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landsc. Ecol.* 21(6):797–807

Holderegger R, Wagner HH. 2006. A brief guide to Landscape Genetics. *Landsc. Ecol.* 21(6):793–96

Hölker F, Wolter C, Perkin EK, Tockner K. 2010. Light pollution as a biodiversity threat. *Trends Ecol. Evol.* 25(12):681–82

Holland JD, Bert DG, Fahrig L. 2004. Determining the Spatial Scale of Species' Response to Habitat. *Bioscience*. 54(3):227–33

Holt RD. 1996. Adaptive Evolution in Source-Sink Environments: Direct and Indirect Effects of Density- Dependence on Niche Evolution. *Oikos*. 75(2):182–92

Holzhauer SIJ, Ekschmitt K, Sander AC, Dauber J, Wolters V. 2006. Effect of historic landscape change on the genetic structure of the bush-cricket Metrioptera roeseli. *Landsc. Ecol.* 21(6):891–99

Hooten MB, Johnson DS, McClintock BT, Morales JM. 2017. *Animal Movement Statistical Models for Telemetry Data*. Boca Raton, Florida, USA: CRC press

Houle M, Fortin D, Dussault C, Courtois R, Ouellet JP. 2010. Cumulative effects of forestry on habitat use by gray wolf (Canis lupus) in the boreal forest. *Landsc. Ecol.* 25(3):419–33

Howeth JG, Mcgaugh SE, Hendrickson DA. 2008. Contrasting demographic and genetic estimates of dispersal in the endangered Coahuilan box turtle: A contemporary approach to conservation. *Mol. Ecol.* 17(19):4209–21

Hubbell SP. 2011. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ, USA: Princeton University Press

Hui FKCC, Warton DI, Foster SD, Dunstan PK. 2013. To mix or not to mix: Comparing the predictive performance of mixture models vs. separate species distribution models. *Ecology*. 94(9):1913–19

Hulme PE. 2008. Contrasting alien and native plant species-area relationships: The importance of spatial grain and extent. *Glob. Ecol. Biogeogr.* 17(5):641–47

Humphreys AM, Govaerts R, Ficinski SZ, Nic Lughadha E, Vorontsova MS. 2019. Global dataset shows geography and life form predict modern plant extinction and rediscovery. *Nat. Ecol. Evol.* 3(7):1043–47

Hussey NE, Kessel ST, Aarestrup K, Cooke SJ, Cowley PD, et al. 2015. Aquatic animal telemetry: A panoramic window into the underwater world. *Science.* 348(6240):1255642

Hutchinson GE. 1957. Concluding Remarks. *Cold Spring Harb. Symp. Quant. Biol.* 22:415–27

Ims RA, Yoccoz NG. 1997. Studying Transfer Processes in Metapopulations: Emigration, Migration, and Colonization. In *Metapopulation Biology: Ecology, Genetics, and Evolution*, eds. IA Hanski, ME Gilpin, pp. 247–66. San Diego, CA,USA: Academic Press

Ives AR. 2015. For testing the significance of regression coefficients, go ahead and logtransform count data. *Methods Ecol. Evol.* 6(7):828–35

James PMA, Coltman DW, Murray BW, Hamelin RC, Sperling FAH. 2011a. Spatial genetic structure of a symbiotic beetle-fungal system: Toward multi-taxa integrated landscape genetics. *PLoS One*. 6(10):e25359

James PMA, Cooke BJ, Brunet BMT, Lumley LM, Sperling FAH, et al. 2015. Life-stage differences in spatial genetic structure in an irruptive forest insect: implications for dispersal and spatial synchrony. *Mol. Ecol.* 24(2):296–309

James PMA, Fleming R a., Fortin M-J. 2010. Identifying significant scale-specific spatial boundaries using wavelets and null models: Spruce budworm defoliation in Ontario, Canada as a case study. *Landsc. Ecol.* 25:873–87

James PMA, Fortin M-J, Fall A, Kneeshaw D, Messier C. 2007. The effects of spatial legacies following shifting management practices and fire on boreal forest age structure. *Ecosystems*. 10(8):1261–77

James PMA, Fortin M-J, Sturtevant BR, Fall A, Kneeshaw D. 2011b. Modelling Spatial Interactions Among Fire, Spruce Budworm, and Logging in the Boreal Forest. *Ecosystems*. 14(1):60–75

James PMA, Huber DPW. 2019. TRIA-Net: 10 years of collaborative research on turning risk into action for the mountain pine beetle epidemic. *Can. J. For. Res.* 49(12):iii–v

Janes JK, Li Y, Keeling CI, Yuen MMSS, Boone CK, et al. 2014. How the mountain pine beetle (Dendroctonus ponderosae) breached the canadian rocky mountains. *Mol. Biol. Evol.* 31(7):1803–15

Janes JK, Worth JRP, Batista PD, Sperling FAH. 2018. Inferring Ancestry and Divergence Events in a Forest Pest Using Low-Density Single-Nucleotide Polymorphisms. *Insect Syst. Divers.* 2(6):

Jay F, Manel S, Alvarez N, Durand EY, Thuiller W, et al. 2012. Forecasting changes in population genetic structure of alpine plants in response to global warming. *Mol. Ecol.* 21(10):2354–68

Johns RC, Bowden JJ, Carleton DR, Cooke BJ, Edwards S, et al. 2019. A conceptual framework for the spruce budworm Early Intervention Strategy: Can outbreaks be stopped? *Forests*. 10(10):1–19

Johnson AF, Lidström S. 2018. The balance between concepts and complexity in ecology. *Nat. Ecol. Evol.* 2(4):585–87

Johnson CJ, Gillingham MP. 2008. Sensitivity of species-distribution models to error, bias, and model design: An application to resource selection functions for woodland caribou. *Ecol. Modell.* 213(2):143–55

Johnson JB, Omland KS. 2004. Model selection in ecology and evolution. *Trends Ecol. Evol.* 19(2):101–8

Jombart T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24(11):1403–5

Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*. 27(21):3070–71

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11(1):94

Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, et al. 2017. Where less may be more: How the rare biosphere pulls ecosystems strings. *ISME J.* 11(4):853–62

Jumentier B, Caye K, Heude B, Lepeule J, François O. 2020. Sparse latent factor regression models for genome-wide and epigenome-wide association studies. *bioRxiv*

Kamm J, Terhorst J, Durbin R, Song YS. 2020. Efficiently Inferring the Demographic History of Many Populations With Allele Count Data. *J. Am. Stat. Assoc.* 115(531):1472–87

Kamvar ZN, Brooks JC, Grünwald NJ. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* 6:208

Kamvar ZN, Tabima JF, Gr̈unwald NJ. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*. 2014(1):1–14

Kareiva P, Wennergren U. 1995. Connecting landscape patterns to ecosystem and population processes. *Nature*. 373(6512):299–302

Kays R, Davidson SC, Berger M, Bohrer G, Fiedler W, et al. 2021. The Movebank system for studying global animal movement and demography. *Methods Ecol. Evol.* 2021(October):1–13

Keeley ATH, Beier P, Keeley BW, Fagan ME. 2017. Habitat suitability is a poor proxy for landscape connectivity during dispersal and mating movements. *Landsc. Urban Plan.* 161:90–102

Keim JL, Lele SR, DeWitt PD, Fitzpatrick JJ, Jenni NS. 2019. Estimating the intensity of use by interacting predators and prey using camera traps. *J. Anim. Ecol.* 88(5):690–701

Keller D, Holderegger R. 2013. Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conserv. Divers.* 6(5):590–97

Keller D, Van Strien MJ, Holderegger R. 2012. Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshw. Biol.* 57(7):1373–84

Kinsley AC, Rossi G, Silk MJ, VanderWaal K. 2020. Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology. *Front. Vet. Sci.* 7:596

Kinziger AP, Hellmair M, McCraney WT, Jacobs DK, Goldsmith G. 2015. Temporal genetic analysis of the endangered tidewater goby: Extinction-colonization dynamics or drift in isolation? *Mol. Ecol.* 24(22):5544–60

Kirby RR, Beaugrand G. 2009. Trophic amplification of climate warming. *Proc. R. Soc. B Biol. Sci.* 276(1676):4095–4103

Kirchheimer B, Schinkel CCF, Dellinger AS, Klatt S, Moser D, et al. 2016. A matter of scale: Apparent niche differentiation of diploid and tetraploid plants may depend on extent and grain of analysis. *J. Biogeogr.* 43(4):716–26

Knapp AK, Smith MD, Hobbie SE, Collins SL, Fahey TJ, et al. 2012. Past, present, and future roles of long-term experiments in the LTER network. *Bioscience*. 62(4):377–89

Knight NL, Vaghefi N, Hansen ZR, Kikkert JR, Pethybridge SJ. 2018. Temporal Genetic Differentiation of Cercospora beticola Populations in New York Table Beet Fields. *Plant*

Dis. 102(11):2074–82

Koen EL, Bowman J, Garroway CJ, Wilson PJ. 2013. The Sensitivity of Genetic Connectivity Measures to Unsampled and Under-Sampled Sites. *PLoS One*. 8(2):

Koen EL, Bowman J, Sadowski C, Walpole AA. 2014. Landscape connectivity for wildlife: Development and validation of multispecies linkage maps. *Methods Ecol. Evol.* 5(7):626–33

Koen EL, Bowman J, Walpole A a. 2012. The effect of cost surface parameterization on landscape resistance estimates. *Mol. Ecol. Resour.* 12(4):686–96

Koen EL, Garroway CJ, Wilson PJ, Bowman J. 2010. The effect of map boundary on estimates of landscape resistance to animal movement. *PLoS One*. 5(7):1–8

Konowalik K, Nosol A. 2021. Evaluation metrics and validation of presence-only species distribution models based on distributional maps with varying coverage. *Sci. Rep.* 11(1):1–15

Kool JT, Moilanen A, Treml EA. 2013. Population connectivity: Recent advances and new perspectives. *Landsc. Ecol.* 28(2):165–85

Korte A, Ashley F. 2013. The advantages and limitations of trait analysis with GWAS : a review Self-fertilisation makes Arabidopsis particularly well suited to GWAS. *Plant Methods*. 9(1):29

Krefting LW. 1974. Moose distribution & habitat selection in north central North America. *Le Nat. Can.* 101:81–100

Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, et al. 2012. Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecol. Lett.* 15(4):378–92

Kuczynski L, Legendre P, Grenouillet G. 2018. Concomitant impacts of climate change, fragmentation and non-native species have led to reorganization of fish communities since the 1980s. *Glob. Ecol. Biogeogr.* 27(2):213–22

Kuemmerlen M, Stoll S, Sundermann A, Haase P. 2016. Long-term monitoring data meet freshwater species distribution models: Lessons from an LTER-site. *Ecol. Indic.* 65:122–32

Lafontaine A, Drapeau P, Fortin D, Gauthier S, Boulanger Y, St-Laurent M-H. 2019. Exposure to historical burn rates shapes the response of boreal caribou to timber harvesting. *Ecosphere*. 10(5):e02739

Lakoff G. 2010. Why it Matters How We Frame the Environment. *Environ. Commun.* 4(1):70–81

Laliberte AS, Ripple WJ. 2004. Range Contractions of North American Carnivores and Ungulates. *Bioscience*. 54(2):123–38

Laliberté J, St-Laurent M-H. 2020. Validation of functional connectivity modeling: The Achilles' heel of landscape connectivity mapping. *Landsc. Urban Plan.* 202(May 2019):103878

Landguth EL, Bearlin A, Day CC, Dunham J. 2017a. CDMetaPOP: an individual-based, eco-evolutionary model for spatially explicit simulation of landscape demogenetics. *Methods Ecol. Evol.* 8(1):4–11

Landguth EL, Bearlin A, Day CC, Dunham J. 2017b. CDMetaPOP: an individual-based, eco-evolutionary model for spatially explicit simulation of landscape demogenetics. *Methods Ecol. Evol.* 8(1):4–11

Landguth EL, Cushman SA, Schwartz MK, McKelvey KS, Murphy MA, Luikart G. 2010. Quantifying the lag time to detect barriers in landscape genetics. *Mol. Ecol.* 19(19):4179–91

Landguth EL, Holden ZA, Mahalovich MF, Cushman SA. 2017c. Using landscape genetics simulations for planting blister rust resistant whitebark pine in the US Northern Rocky Mountains. *Front. Genet.* 8:9

Larroque J, Legault S, Johns R, Lumley L, Cusson M, et al. 2019. Temporal variation in spatial genetic structure during population outbreaks: Distinguishing among different potential drivers of spatial synchrony. *Evol. Appl.* 12(10):1931–45

Larroque J, Ruette S, Vandel JM, Devillard S. 2016. Divergent landscape effects on genetic differentiation in two populations of the European pine marten (Martes martes). *Landsc. Ecol.* 31(3):517–31

Larroque J, Wittische J, James PMA. 2022. Quantifying and predicting population connectivity of an outbreaking forest insect pest. *Landsc. Ecol.* 37(1):

Latham ADM, Latham MC, Boyce MS. 2011a. Habitat selection and spatial relationships of black bears (*Ursus americanus*) with woodland caribou (*Rangifer tarandus caribou*) in northeastern Alberta. *Can. J. Zool.* 89(4):267–77

Latham ADM, Latham MC, Boyce MS, Boutin S. 2011b. Movement responses by wolves to industrial linear features and their effect on woodland caribou in northeastern alberta. *Ecol.*

Appl. 21(8):2854–65

Latham ADM, Latham MC, Boyce MS, Boutin S. 2013a. Spatial relationships of sympatric wolves (Canis lupus) and coyotes (C. latrans) with woodland caribou (Rangifer tarandus caribou) during the calving season in a human-modified boreal landscape. *Wildl. Res.* 40(3):250–60

Latham ADM, Latham MC, Boyce MS, Boutin S. 2013b. Spatial relationships of sympatric wolves (Canis lupus) and coyotes (C. latrans) with woodland caribou (Rangifer tarandus caribou) during the calving season in a human-modified boreal landscape. *Wildl. Res.* 40(3):250–60

Latham ADM, Latham MC, McCutchen NA, Boutin S. 2011c. Invading white-tailed deer change wolf-caribou dynamics in northeastern Alberta. *J. Wildl. Manage.* 75(1):204–12

Lauterjung MB, Montagna T, Bernardi AP, da Silva JZ, da Costa NCF, et al. 2019. Temporal changes in population genetics of six threatened Brazilian plant species in a fragmented landscape. *For. Ecol. Manage.* 435:144–50

Lawson Handley LJ, Estoup A, Evans DM, Thomas CE, Lombaert E, et al. 2011. Ecological genetics of invasive alien species. *BioControl*. 56(4):409–28

Leblois R, Pudlo P, Néron J, Bertaux F, Reddy Beeravolu C, et al. 2014. Maximum-likelihood inference of population size contractions from microsatellite data. *Mol. Biol. Evol.* 31(10):2805–23

Leblond M, Dussault C, Ouellet JP. 2013. Avoidance of roads by large herbivores and its relation to disturbance intensity. *J. Zool.* 289(1):32–40

Leblond M, Dussault C, Ouellet JP, St-Laurent MH. 2016. Caribou avoiding wolves face increased predation by bears – Caught between Scylla and Charybdis. *J. Appl. Ecol.* 53(4):1078–87

Leblond M, Frair J, Fortin D, Dussault C, Ouellet JP, Courtois R. 2011. Assessing the influence of resource covariates at multiple spatial scales: An application to forest-dwelling caribou faced with intensive human activity. *Landsc. Ecol.* 26(10):1433–46

Legendre P. 2019. A temporal beta-diversity index to identify sites that have changed in exceptional ways in space–time surveys. *Ecol. Evol.* 9(6):3500–3514

Legendre P, Borcard D, Roberts DW. 2012. Variation partitioning involving orthogonal spatial eigenfunction submodels. *Ecology*. 93(5):1234–40

Legendre P, De Cáceres M. 2013. Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecol. Lett.* 16(8):951–63

Legendre P, Fortin M-J. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol. Ecol. Resour.* 10(5):831–44

Legendre P, Fortin M-J, Borcard D. 2015. Should the Mantel test be used in spatial analysis? *Methods Ecol. Evol.* 6(11):1239–47

Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia*. 129(2):271–80

Legendre P, Legendre L. 2012. *Numerical Ecology*. Amsterdam, Netherlands: Elsevier. 3rd

Englis ed.

Legge S, Lindenmayer DB, Robinson NM, Scheele BC, Southwell DM, Wintle lrendan C. 2018. *Monitoring Threatened Species and Ecological Communities*. Clayton (Australia): CSIRO Publishing

Legrand D, Baguette M, Dubois Q, Turlure C, Schtickzelle N. 2021. Congruent Genetic and Demographic Dispersal Rates in a Natural Metapopulation at Equilibrium. *Genes.* 12(3):362

Leigh DM, Hendry AP, Vázquez-Domínguez E, Friesen VL. 2019. Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. *Evol. Appl.* 12(8):1505–12

Leitão RP, Zuanon J, Villéger S, Williams SE, Baraloto C, et al. 2016. Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. R. Soc. B Biol. Sci.* 283(1828):20160084

Leroux SJ, Schmiegelow FKA, Cumming SG, Lessard RB, Nagy J. 2007. Accounting for system dynamics in rerserve design. *Ecol. Appl.* 17(7):1954–66

Lesmerises F, Dussault C, St-Laurent M-H. 2012. Wolf habitat selection is shaped by human activities in a highly managed boreal forest. *For. Ecol. Manage.* 276:125–31

Lesmerises R, Ouellet JP, Dussault C, St-Laurent M-H. 2013. The influence of landscape matrix on isolated patch use by wide-ranging animals: Conservation lessons for woodland caribou. *Ecol. Evol.* 3(9):2880–91

Levin SA. 1992. The Problem of Pattern and Scale in Ecology. *Ecology*. 73(6):1943–67

Levy S. 2012. Rise of the coyote: the new top dog. *Nature*. 485(7398):296–97

Lewis CW, Hodges KE, Koehler GM, Mills LS. 2011. Influence of stand and landscape features on snowshoe hare abundance in fragmented forests. *J. Mammal.* 92(3):561–67

Lewis KP, Gullage SE, Fifield DA, Jennings DH, Mahoney SP. 2017. Manipulations of black bear and coyote affect caribou calf survival. *J. Wildl. Manage.* 81(1):122–32

Liberg O, Chapron G, Wabakken P, Pedersen HC, Thompson Hobbs N, Sand H. 2012. Shoot, shovel and shut up: Cryptic poaching slows restoration of a large carnivore in Europe. *Proc. R. Soc. B Biol. Sci.* 279(1730):910–15

Linck E, Battey CJ. 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol. Ecol. Resour.* 19(3):639–47

Lindgren FK, Rue H, Lindström J, Kent JT, Diggle PJ, et al. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields : the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B*. 73(4):423–98

Logan JA, Amman GD. 1991. Temperature-dependent development of the mountain pine beetle (coleoptera: Scolytidae) and simulation of its phenology. *Can. Entomol.* 123(5):1083–94

Logan JA, Bentz BJ. 1999. Model Analysis of Mountain Pine Beetle (Coleoptera: Scolytidae) Seasonality. *Environ. Entomol.* 28(6):924–34

Logan JA, Powell JA. 2001. Ghost Forests, Global Warming, and the Mountain Pine Beetle (Coleoptera: Scolytidae). *Am. Entomol.* 47(3):160–72

Loredo I, Van Vuren D, Morrison ML. 1996. Habitat use and migration behavior of the California tiger salamander. *J. Herpetol.* 30(2):282–85

Lotterhos KE, Whitlock MC. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol. Ecol.* 23(9):2178–92

Low-Décarie E, Chivers C, Granados M. 2014. Rising complexity and falling explanatory power in ecology. *Front. Ecol. Environ.* 12(7):412–18

Lowe WH, Allendorf FW. 2010. What can genetics tell us about population connectivity? *Mol. Ecol.* 19(15):3038–51

Lowe WH, Kovach RP, Allendorf FW. 2017. Population Genetics and Demography Unite Ecology and Evolution. *Trends Ecol. Evol.* 32(2):141–52

Lundgren E, Ralph PL. 2019. Are populations like a circuit? Comparing isolation by resistance to a new coalescent-based method. *Mol. Ecol. Resour.* 19(6):1388–1406

Lynch MDJ, Neufeld JD. 2015. Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13(4):217–29

MacArthur RH, Wilson EO. 1967. *The Theory of Island Biogeography*. Princeton, NJ, USA: Princeton University Press

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz F a. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecol. Appl.* 10(3):689–710

Maguire KC, Nieto-Lugilde D, Blois JL, Fitzpatrick MC, Williams JW, et al. 2016. Controlled comparison of species- and community-level models across novel climates and communities. *Proc. R. Philos. Soc. - B*. 283(1826):20152817

Manel S, Holderegger R. 2013. Ten years of landscape genetics. *Trends Ecol. Evol.* 28(10):614–21

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: Combining landscape ecology and population genetics. *Trends Ecol. Evol.* 18(4):189–97

Månsson J, Andrén H, Pehrson Å, Bergström R. 2007. Moose browsing and forage availability: a scale-dependent relationship? *Can. J. Zool.* 85(3):372–80

Manzoor SA, Griffiths G, Lukac M. 2018. Species distribution model transferability and model grain size-finer may not always be better. *Sci. Rep.* 8(1):1–9

Maréchaux I, Rodrigues ASL, Charpentier A. 2017. The value of coarse species range maps to inform local biodiversity conservation in a global context. *Ecography.* 40(10):1166–76

Marrotte RR, Bowman J, Brown MGC, Cordes C, Morris KY, et al. 2017. Multi-species genetic connectivity in a terrestrial habitat network. *Mov. Ecol.* 5(1):1–11

Marth GT, Czabarka E, Murvai J, Sherry ST. 2004. The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations. *Genetics*. 166(1):351–72

Martin AE, Fahrig L. 2012. Measuring and selecting scales of effect for landscape predictors in species-habitat models. *Ecol. Appl.* 22(8):2277–92

Martin AE, Fahrig L. 2016. Reconciling contradictory relationships between mobility and extinction risk in human-altered landscapes. *Funct. Ecol.* 30(9):1558–67

Martins TG, Simpson D, Lindgren FK, Rue H. 2013. Bayesian computing with INLA: New features. *Comput. Stat. Data Anal.* 67:68–83

Maruyama T, Fuerstt PA. 1985. Population bottlenecks and nonequilibrium models in opulation genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*. 111(3):675–89

Masson-Delmotte V, Zhai P, Pirani A, Connors LS, Péan C, et al. 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press

Mathew B, Holand AM, Koistinen P, Léon J, Sillanpää MJ. 2016. Reparametrization-based estimation of genetic parameters in multi-trait animal model using Integrated Nested Laplace Approximation. *Theor. Appl. Genet.* 129(2):215–25

Maxwell SL, Fuller RA, Brooks TM, Watson JEM. 2016. Biodiversity: The ravages of guns, nets and bulldozers. *Nature*. 536(7615):143–45

May RM. 1974. Biological Populations with Nonoverlapping Generations : Stable Points , Stable Cycles , and Chaos. *Science.* 186(4164):645–47

Maynard AJ, Ambrose L, Cooper RD, Chow WK, Davis JB, et al. 2017. Tiger on the prowl: Invasion history and spatio-temporal genetic structure of the Asian tiger mosquito Aedes albopictus (Skuse 1894) in the Indo-Pacific. *PLoS Negl. Trop. Dis.* 11(4):1–27

Mayor SJ, Schneider DC, Schaefer JA, Mahoney SP. 2009. Habitat selection at multiple scales. *Ecoscience*. 16(2):238–47

Mayrand P, Filotas E, Wittische J, James PMA. 2019a. The role of dispersal, selection, and timing of sampling on the false discovery rate of loci under selection during geographic range expansion. *Genome*. 62(11):715–27

Mayrand P, Filotas É, Wittische J, James PMA. 2019b. The role of dispersal, selection, and timing of sampling on the false discovery rate of loci under selection during geographic range expansion. *Genome*. 62(11):715–27

McCambridge WF. 1971. Temperature Limits of Flight of the Mountain Pine Beetle, Dendroctonus ponderosae. *Ann. Entomol. Soc. Am.* 64(2):534–35

McCauley DJ, Pinsky ML, Palumbi SR, Estes JA, Joyce FH, Warner RR. 2015. Marine defaunation: Animal loss in the global ocean. *Science.* 347(6219):247–54

McClure ML, Hansen AJ, Inman RM. 2016. Connecting models to movements: testing connectivity model predictions against empirical migration and dispersal data. *Landsc. Ecol.* 31(7):1419–32

McInturff A, Xu W, Wilkinson CE, Dejid N, Brashares JS. 2020. Fence Ecology: Frameworks for Understanding the Ecological Effects of Fences. *Bioscience*. 70(11):971–85

McKenzie HW, Merrill EH, Spiteri RJ, Lewis MA. 2012. How linear features alter predator movement and the functional response. *Interface Focus*. 2(2):205–16

McLoughlin PD, Dzus E, Wynes B, Boutin S. 2003. Declines in populations of woodland caribou. *J. Wildl. Manage.* 67(4):755–61

McRae BH. 2006. Isolation By Resistance. *Evolution.* 60(8):1551

McRae BH, Dickson BG, Keitt TH, Shah VB. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*. 89(10):2712–24

Meirmans PG. 2012. The trouble with isolation by distance. *Mol. Ecol.* 21(12):2839–46

Merckx B, Steyaert M, Vanreusel A, Vincx M, Vanaverbeke J. 2011. Null models reveal preferential sampling, spatial autocorrelation and overfitting in habitat suitability modelling. *Ecol. Modell.* 222(3):588–97

Michelot T, Blackwell PG, Chamaillé-Jammes S, Matthiopoulos J. 2020. Inference in MCMC step selection models. *Biometrics*. 76(2):438–47

Milner-Gulland EJ, Shea K. 2017. Embracing uncertainty in applied ecology. *J. Appl. Ecol.* 54(6):2063–68

Miranda-Soberanis VF, Yee TW. 2019. New Link Functions for Distribution-Specific Quantile Regression Based on Vector Generalized Linear and Additive Models. *J. Probab. Stat.* 1:3493628

Mock KE, Bentz BJ, O'Neill EM, Chong JP, Orwin J, Pfrender ME. 2007. Landscape-scale genetic variation in a forest outbreak species, the mountain pine beetle (Dendroctonus ponderosae). *Mol. Ecol.* 16:553–68

Moilanen A, Wilson K, Possingham H. 2009. *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools*. Oxford, United Kingdom: Oxford University Press

Moon JB, Dewitt TH, Errend MN, Bruins RJF, Kentula ME, et al. 2017. Model application niche analysis: assessing the transferability and generalizability of ecological models. *Ecosphere*. 8(10):e01974

Moore JS, Harris LN, Le Luyer J, Sutherland BJG, Rougemont Q, et al. 2017. Genomics and telemetry suggest a role for migration harshness in determining overwintering habitat choice, but not gene flow, in anadromous Arctic Char. *Mol. Ecol.* 26(24):6784–6800

Moraes AM, Ruiz-Miranda CR, Ribeiro MC, Grativol AD, da S. Carvalho C, et al. 2017. Temporal genetic dynamics of reintroduced and translocated populations of the endangered golden lion tamarin (Leontopithecus rosalia). *Conserv. Genet.* 18(5):995–1009

Morales JM, Ellner SP. 2002. Scaling up animal movements in heterogeneous landscapes: The importance of behavior. *Ecology*. 83(8):2240–47

Morin RS, Liebhold AM, Gottschalk KW. 2009. Anisotropic spread of hemlock woolly adelgid in the eastern United States. *Biol. Invasions*. 11(10):2341–50

Mouillot D, Bellwood DR, Baraloto C, Chave J, Galzin R, et al. 2013. Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. *PLoS Biol.* 11(5):e1001569

Mountain Pine Beetle Population Forecast Survey. 2017. Alberta Agriculture and Forestry. https://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/formain15808/$FILE/RvalueMap_2016_11x17.pdf

Muñoz-Mas R, Gil-Martínez E, Oliva-Paterna FJ, Belda EJ, Martínez-Capel F. 2019. Tree-based ensembles unveil the microhabitat suitability for the invasive bleak (Alburnus alburnus L.) and pumpkinseed (Lepomis gibbosus L.): Introducing XGBoost to eco-informatics. *Ecol. Inform.* 53:100974

Murdoch A, Mantyka-Pringle C, Sharma S. 2020. The interactive effects of climate change and land use on boreal stream fish communities. *Sci. Total Environ.* 700(5):1345418

Murphy MA, Dezzani R, Pilliod DS, Storfer AT. 2010. Landscape genetics of high mountain frog metapopulations. *Mol. Ecol.* 19(17):3634–49

Murphy MA, Evans JS, Cushman SA, Storfer AT. 2008. Representing genetic variation as continuous surfaces: An approach for identifying spatial dependency in landscape genetic studies. *Ecography.* 31(6):685–97

Murray GGR, Wang F, Harrison EM, Paterson GK, Mather AE, et al. 2016. The effect of genetic structure on molecular dating and tests for temporal signal. *Methods Ecol. Evol.* 7(1):80–89

Musetta-Lambert JL, Enanga EM, Teichert S, Creed IF, Kidd KA, et al. 2019. Industrial innovation and infrastructure as drivers of change in the Canadian boreal zone. *Environ. Rev.* 27(3):275–94

Naeem S. 2002. Ecosystem consequences of biodiversity loss: The evolution of a paradigm. *Ecology*. 83(6):1537–52

Naujokaitis-Lewis IR, Rico Y, Lovell J, Fortin MJ, Murphy MA. 2013. Implications of incomplete networks on estimation of landscape genetic connectivity. *Conserv. Genet.* 14(2):287–98

Neilson EW, Avgar T, Burton AC, Broadley K, Boutin S. 2018a. Animal movement affects interpretation of occupancy models from camera-trap surveys of unmarked animals. *Ecosphere*. 9(1):e02092

Neilson EW, Avgar T, Cole Burton A, Broadley K, Boutin S. 2018b. Animal movement affects interpretation of occupancy models from camera-trap surveys of unmarked animals. *Ecosphere*. 9(1):

Nelson G. 1978. From Candolle to croizat: Comments on the history of biogeography. *J. Hist. Biol.* 11(2):269–305

Newbury TL, Simon NPP. 2005. The effects of clearcutting on snowshoe hare (Lepus americanus) relative abundance in central Labrador. *For. Ecol. Manage.* 210(1–3):131–42

Newman EA, Kennedy MC, Falk DA, McKenzie D. 2019. Scaling and complexity in landscape ecology. *Front. Ecol. Evol.* 7:293

Nicholson E, Possingham HP. 2007. Making conservation decisions under uncertainty for the persistence of multiple species. *Ecol. Appl.* 17(1):251–65

Nicolau PG, Sørbye SH, Yoccoz NG. 2020. Incorporating capture heterogeneity in the estimation of autoregressive coefficients of animal population dynamics using capture–recapture data. *Ecol. Evol.* 10(23):12710–26

Niku J, Hui FKC, Taskinen S, Warton DI. 2019. gllvm: Fast analysis of multivariate abundance data with generalized linear latent variable models in r. *Methods Ecol. Evol.* 10(12):2173–82

Nustad HE, Page CM, Reiner AH, Zucknick M, Leblanc M. 2018. A Bayesian mixed modeling approach for estimating heritability. *BMC Proc.* 12(Suppl.(31):117–22

O'Connor TD, Fu W, Turner E, Mychaleckyj JC, Logsdon B, et al. 2015. Rare variation facilitates inferences of fine-scale population structure in humans. *Mol. Biol. Evol.* 32(3):653–60

Odum EP. 1950. Bird Populations of the Highlands (North Carolina) Plateau in Relation to Plant Succession and Avian Invasion. *Ecology*. 31(4):587–605

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, et al. 2018. vegan: community ecology package. R package version 2.5-2. https://CRAN.R-project.org/package=vegan

Ontiveros VJ, Capitán JA, Casamayor EO, Alonso D. 2021. The characteristic time of ecological communities. *Ecology*. 102(2):e03247

Opitz T. 2017. Latent Gaussian modeling and INLA: A review with focus on space-time applications. *J. la Société Franaise Stat.* 158(3):62–85

Opitz T, Huser R, Bakka H, Rue H. 2018. INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *Extremes*. 21(3):441–62

Osborne JL, Loxdale HD, Woiwod IP. 2002. Monitoring insect dispersal methods and approaches. *Dispersal Ecol. 42nd Symp. Br. Ecol. Soc.*, pp. 24–49. Reading, UK: Blackwell Publishing

Ovaskainen O, Tikhonov G, Dunson D, Grøtan V, Engen S, et al. 2017a. How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proc. R. Soc. B Biol. Sci.* 284(1855):1–7

Ovaskainen O, Tikhonov G, Norberg A, Guillaume Blanchet F, Duan L, et al. 2017b. How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecol. Lett.* 20(5):561–76

Oxley DJ, Fenton MB, Carmody GR. 1974. The effects of roads on populations of small mammals. *J. Appl. Ecol.* 11(1):51–59

Palm EC, Fluker S, Nesbitt HK, Jacob AL, Hebblewhite M. 2020. The long road to protecting critical habitat for species at risk: The case of southern mountain woodland caribou. *Conserv. Sci. Pract.* 2(7):1–13

Paradis E. 2010. Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*. 26(3):419–20

Parrott L. 2002. Complexity and the limits of ecological engineering. *Trans. Am. Soc. Agric. Eng.* 45(5):1697–1702

Pavlacky DC, Goldizen AW, Prentis PJ, Nicholls JA, Lowe AJ. 2009. A landscape genetics approach for quantifying the relative influence of historic and contemporary habitat heterogeneity on the genetic connectivity of a rainforest bird. *Mol. Ecol.* 18(14):2945–60

Paz-Vinas I, Loot G, Stevens VM, Blanchet S. 2015. Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Mol. Ecol.* 24(18):4586–4604

Paz-Vinas I, Quéméré E, Chikhi L, Loot G, Blanchet S. 2013. The demographic history of populations experiencing asymmetric gene flow: Combining simulated and empirical data.

Mol. Ecol. 22(12):3279–91

Peng J, Zhou Y, Wang K. 2021. Multiplex gene and phenotype network to characterize shared genetic pathways of epilepsy and autism. *Sci. Rep.* 11(1):1–16

Pereira HM, Leadley PW, Proença V, Alkemade R, Scharlemann JPW, et al. 2010. Scenarios for Global Biodiversity in the 21st Century. *Science.* 330(6010):1496–1501

Pereira P, Teixeira J, Velo-Antón G. 2018. Allele surfing shaped the genetic structure of the European pond turtle via colonization and population expansion across the Iberian Peninsula from Africa. *J. Biogeogr.* 45(9):2202–15

Peres-Neto PR, Legendre P. 2010. Estimating and controlling for spatial structure in the study of ecological communities. *Glob. Ecol. Biogeogr.* 19(2):174–84

Peres-Neto PR, Legendre P, Dray S, Borcard D. 2006. Variation partitioning of species data matrices : estimation and comparison of fractions. *Ecology*. 87(10):2614–25

Pérez-Portela R, Wangensteen OS, Garcia-Cisneros A, Valero-Jiménez C, Palacín C, Turon X. 2019. Spatio-temporal patterns of genetic variation in Arbacia lixula, a thermophilous sea urchin in expansion in the Mediterranean. *Heredity.* 122(2):244–59

Peterman WE. 2018. ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods Ecol. Evol.* 9(6):1638–47

Peterman WE, Connette GM, Semlitsch RD, Eggert LS. 2014. Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol. Ecol.* 23(10):2402–13

Peterman WE, Pope NS. 2021. The use and misuse of regression models in landscape genetic analyses. *Mol. Ecol.* 30(1):37–47

Peters DPC, Havstad KM, Cushing J, Tweedie C, Fuentes O, Villanueva-Rosales N. 2014. Harnessing the power of big data: Infusing the scientific method with machine learning to transform ecology. *Ecosphere*. 5(6):1–15

Peterson AT, Navarro-Sigüenza AG, Gordillo A. 2018. Assumption-versus data-based approaches to summarizing species' ranges. *Conserv. Biol.* 32(3):568–75

Petitpierre B, Broennimann O, Kueffer C, Daehler C, Guisan A. 2017. Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Glob. Ecol. Biogeogr.* 26(3):275–87

Phillis CC, O'Regan SM, Green SJ, Bruce JEB, Anderson SC, et al. 2013. Multiple pathways to conservation success. *Conserv. Lett.* 6(2):98–106

Pickell PD, Andison DW, Coops NC, Gergel SE, Marshall PL. 2015. The spatial patterns of anthropogenic disturbance in the western Canadian boreal forest following oil and gas development. *Can. J. For. Res.* 45(6):732–43

Pickett STA, Cadenasso ML. 1995. Landscape Ecology: Spatial Heterogeneity in Ecological Systems. *Science.* 269(5222):331–34

Pilosof S, Porter MA, Pascual M, Kéfi S. 2017. The multilayer nature of ecological networks. *Nat. Ecol. Evol.* 1(4):

Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, et al. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science.* 344(6187):987–97

Pither R, Brien PO, Brennan A, Hirsh-pearson K. 2021. Areas Important for Ecological Connectivity Throughout Canada. *bioRxiv*. 1–23

Plant E, King R, Kath J. 2021. Statistical comparison of additive regression tree methods on ecological grassland data. *Ecol. Inform.* 61(7):101198

Poff NLR, Larson EI, Salerno PE, Morton SG, Kondratieff BC, et al. 2018. Extreme streams: species persistence and genomic change in montane insect populations across a flooding gradient. *Ecol. Lett.* 21(4):525–35

Polfus JL, Hebblewhite M, Heinemeyer K. 2011. Identifying indirect habitat loss and avoidance of human infrastructure by northern mountain woodland caribou. *Biol. Conserv.* 144(11):2637–46

Polychronopoulos D, King JWD, Nash AJ, Tan G, Lenhard B. 2017. Conserved non-coding elements: Developmental gene regulation meets genome organization. *Nucleic Acids Res.* 45(22):12611–24

Ponisio LC, de Valpine P, Michaud N, Turek D. 2020. One size does not fit all: Customizing MCMC methods for hierarchical models using NIMBLE. *Ecol. Evol.* 10(5):2385–2416

Poor EE, Loucks C, Jakes A, Urban DL. 2012. Comparing Habitat Suitability and Connectivity Modeling Methods for Conserving Pronghorn Migrations. *PLoS One*. 7(11):

Poorter L, Craven D, Jakovac CC, Sande MT Van Der, Amissah L, et al. 2021. Multidimensional tropical forest recovery. . (December):1–8

Poulsen JR, Koerner SE, Moore S, Medjibe VP, Blake S, et al. 2017. Poaching empties critical Central African wilderness of forest elephants. *Curr. Biol.* 27(4):134–35

Powers RP, Jetz W. 2019. Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. *Nat. Clim. Chang.* 9(4):323–29

Preston FW. 1960. Time and space and the variation of species. *Ecology*. 41(4):611–27

Privé F, Aschard H, Blum MGB. 2019. Efficient Implementation of Penalized Regression for Genetic Risk Prediction. *Genetics*. 212(1):65–74

Privé F, Luu K, Vilhjàlmsson BJ, Blum MGB. 2020. Performing Highly Efficient Genome Scans for Local Adaptation with R Package pcadapt Version 4. *Mol. Biol. Evol.* 37(7):2153–54

Pulliam HR. 1988. Source, Sinks, and Population Regulation. *Am. Nat.* 132(5):652–61

Qian H. 2009. Beta diversity in relation to dispersal ability for vascular plants in North America. *Glob. Ecol. Biogeogr.* 18(3):327–32

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria.

Raffa KF, Aukema BH, Bentz BJ, Carroll AL, Hicke J a., et al. 2008. Cross-scale Drivers of Natural Disturbances Prone to Anthropogenic Amplification: The Dynamics of Bark Beetle Eruptions. *Bioscience*. 58(6):501

Rands MRW, Adams WM, Bennun L, Butchart SHM, Clements A, et al. 2010. Biodiversity

Conservation: Challenges Beyond 2010. *Science.* 329(5997):1298–1303

Rasmussen CE, Williams CKI. 2006. *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: The Massachusetts Institute of Technology Press

Rayfield B, Fortin M-J, Fall A. 2010. The sensitivity of least-cost habitat graphs to relative cost surface values. *Landsc. Ecol.* 25(4):519–32

Redford KH. 1992. The Empty Forest. *Bioscience*. 42(6):412–22

Reid BN, Thiel RP, Palsbøll PJ, Peery MZ. 2016. Linking genetic kinship and demographic analyses to characterize dispersal: Methods and application to blanding's turtle. *J. Hered.* 107(7):603–14

Rejmánek M, Richardson DM. 1996. What attributes make some plant species more invasive? *Ecology*. 77(6):1655–61

Rellstab C, Dauphin B, Exposito-Alonso M. 2021. Prospects and limitations of genomic offset in conservation management. *Evol. Appl.* 14(5):1202–12

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R. 2015. A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.* 24(17):4348–70

Rempel RS, Elkie PC, Rodgers AR, Gluck MJ. 1997. Timber-management and natural-disturbance effects on moose habitat : landscape evaluation. *J. Wildl. Manage.* 61(2):517–24

Rencher AC, Pun FC. 1980. Inflation of R2 in Best Subset Regression. *Technometrics*. 22(1):49–53

Richardson JL, Brady SP, Wang IJ, Spear SF. 2016. Navigating the pitfalls and promise of landscape genetics. *Mol. Ecol.* 25(4):849–63

Ripple WJ, Wolf C, Newsome TM, Betts MG, Ceballos G, et al. 2019. Are we eating the world's megafauna to extinction? *Conserv. Lett.* 12(3):e12627

Robledo-Arnuncio JJ, Rousset F. 2010. Isolation by distance in a continuous population under stochastic demographic fluctuations. *J. Evol. Biol.* 23(1):53–71

Rockström J, W. Steffen, K. Noone, Å. Persson, F. S. Chapin, et al. 2009. A safe operating space for humanity. *Nature*. 461:472–75

Rogers JS. 1972. Measures of genetic similarity and genetic distances. In *Studies in Genetics VII*, ed. MR Wheeler, pp. 145–53. Austin: The University of Texas

Rosa L, Davis KF, Rulli MC, D'Odorico P. 2017. Environmental consequences of oil production from oil sands. *Earth's Futur.* 5(2):158–70

Rousset F. 1997. Genetic Differentiation and Estimation of Gene Flow from FStatistics Under Isolation by Distance. *Genetics*. 145(4):1219–28

Rousset F, Ferdy J-B. 2014. Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography.* 37(8):781–90

Row JR, Knick ST, Oyler-McCance SJ, Lougheed SC, Fedy BC. 2017. Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecol. Evol.* 7(11):3751–61

RStudio Team. 2018. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Rubin BER, Jones BM, Hunt BG, Kocher SD. 2019. Rate variation in the evolution of non-coding DNA associated with social evolution in bees. *Philos. Trans. R. Soc. B Biol. Sci.* 374(1777):1–13

Rudnick D, Ryan MG, Harmon ME, Birdsey R a, Giardina CP, et al. 2012. The Role of Landscape Connectivity in Planning and Implementating Conservation and Restoration Priorities. *Issue Ecol.* 13(16):1–16

Rudolph TD, Drapeau P, Imbeau L, Brodeur V, Légaré S, St-Laurent M-H. 2017. Demographic responses of boreal caribou to cumulative disturbances highlight elasticity of range-specific tolerance thresholds. *Biodivers. Conserv.* 26(5):1179–98

Rue H, Martino S, Chopin N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 71(2):319–92

Rue H, Riebler A, Sørbye SH, Illian JB, Simpson DP, Lindgren FK. 2017. Bayesian computing with INLA: A review. *Annu. Rev. Stat. Its Appl.* 4:395–421

Safranyik L, Carroll AL. 2006. The biology and epidemiology of the mountain pine beetle in lodgepole pine forests. In *The Mountain Pine Beetle: A Synthesis of Its Biology,* *Management and Impacts on Lodgepole Pine.*, eds. L Safranyik, B Wilson, pp. 3–66. Victoria, British Columbia, Canada: Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre

Safranyik L, Carroll AL, Régnière J, Langor DW, Riel WG, et al. 2010a. Potential for range expansion of mountain pine beetle into the boreal forest of North America. *Can. Entomol.* 142(5):415–42

Safranyik L, Carroll AL, Régnière J, Langor DW, Riel WG, et al. 2010b. Potential for range expansion of mountain pine beetle into the boreal forest of North America. *Can. Entomol.* 142(5):415–42

Safranyik L, Llnton DA. 1998. Mortality of mountain pine beetle larvae, Dendroctonus ponderosae (Coleoptera: Scolytidae) in logs of lodgepole pine (Pinus contorta var. latifolia) at constant low temperatures. *J. Entomol. Soc. Br. Columbia*. 95(0):81-88–88

Safranyik L, Shrimpton DM, Whitney HS. 1975. An interpretation of the interaction between lodgepole pine, the mountain pine beetle, and its associated blue stain fungi in western Canada. *Manag. Lodg. Pine Ecosyst. Symp. Proc.*, pp. 406–28. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0

Safranyik L, Wilson B. 2006. *The Mountain Pine Beetle : A Synthesis of Biology, Management, and Impacts on Lodgepole Pine*

Samarasekera GDNG, Bartell N V, Lindgren BS, Cooke JEK, Davis CS, et al. 2012. Spatial genetic structure of the mountain pine beetle (Dendroctonus ponderosae) outbreak in western Canada: historical patterns and contemporary dispersal. *Mol. Ecol.* 21(12):2931–48

Sambaraju KR, Carroll AL, Zhu J, Stahl K, Moore RD, Aukema BH. 2012. Climate change could alter the distribution of mountain pine beetle outbreaks in western Canada. *Ecography.* 35:211–23

Sawitri, Tani N, Na'iem M, Widiyatno, Indrioko S, et al. 2020. Potential of Genome-Wide association studies and Genomic Selection to improve productivity and quality of commercial timber species in tropical rainforest, a case study of Shorea platyclados. *Forests*. 11(2):239

Schraiber JG, Akey JM. 2015. Methods and models for unravelling human evolutionary history. *Nat. Rev. Genet.* 16(12):727–40

Schrey AW, Liebl AL, Richards CL, Martin LB. 2014. Range expansion of house sparrows (Passer domesticus) in kenya: Evidence of genetic admixture and human-mediated dispersal. *J. Hered.* 105(1):60–69

Scrucca L. 2013. GA : A Package for Genetic Algorithms in R. *J. Stat. Softw.* 53(4):1–37

Scrucca L. 2017. On some extensions to GA package: hybrid optimisation, parallelisation and islands evolution. *R J.* 9(1):187–206

Segelbacher G, Cushman SA, Epperson BK, Fortin M-J, Francois O, et al. 2010. Applications of landscape genetics in conservation biology: Concepts and challenges. *Conserv. Genet.* 11(2):375–85

Segura-García I, Garavelli L, Tringali M, Matthews T, Chérubin LM, et al. 2019. Reconstruction of larval origins based on genetic relatedness and biophysical modeling. *Sci. Rep.* 9(1):1–9

Selle ML, Steinsland I, Hickey JM, Gorjanc G. 2019. Flexible modelling of spatial variation in agricultural field trials with the R package INLA. *Theor. Appl. Genet.* 132(12):3277–93

Serrouya R, Dickie M, Lamb C, van Oort H, Kelly AP, et al. 2021. Trophic consequences of terrestrial eutrophication for a threatened ungulate. *Proc. R. Soc. B Biol. Sci.* 288(1943):20202811

Serrouya R, Seip DR, Hervieux D, McLellan BN, McNay RS, et al. 2019. Saving endangered species using adaptive management. *Proc. Natl. Acad. Sci. U. S. A.* 116(13):6181–86

Shegelski VA, Campbell EO, Thompson KM, Whitehouse CM, Sperling FAH. 2021. Source and spread dynamics of mountain pine beetle in central Alberta, Canada. *Can. Entomol.* 153(3):314–26

Shirk AJ, Landguth EL, Cushman SA. 2018. A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Mol. Ecol. Resour.* 18(1):55–67

Shirk AJ, Landguth EL, Cushman SA. 2021. The effect of gene flow from unsampled demes in landscape genetic analysis. *Mol. Ecol. Resour.* 21(2):394–403

Shokralla S, Spall JL, Gibson JF, Hajibabaei M. 2012. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* 21(8):1794–1805

Shore TL, Brooks JE, Stone JE. 2004. Mountain Pine Beetle Symposium: Challenges and Solutions. Kelowna, British Columbia, Canada

Sidak Z. 1967. Rectangular confidence regions for the means of multivariate normal distributions. *J. Am. Stat. Assoc.* 62(318):626–33

Simberloff D, Martin JL, Genovesi P, Maris V, Wardle DA, et al. 2013. Impacts of biological invasions: What's what and the way forward. *Trends Ecol. Evol.* 28(1):58–66

Sinclair SP, Milner-Gulland EJ, Smith RJ, McIntosh EJ, Possingham HP, et al. 2018. The use, and usefulness, of spatial conservation prioritizations. *Conserv. Lett.* 11(6):1–7

Skoglund P, Sjödin P, Skoglund T, Lascoux M, Jakobsson M. 2014. Investigating population history using temporal genetic differentiation. *Mol. Biol. Evol.* 31(9):2516–27

Slatkin M. 2005. Seeing ghosts: The effect of unsampled populations on migration rates estimated for sampled populations. *Mol. Ecol.* 14(1):67–73

Slatkin M, Excoffier L. 2012. Serial founder effects during range expansion: A spatial analog of genetic drift. *Genetics*. 191(1):171–81

Smith KG, Ficht EJ, Hobson D, Sorensen TC, Hervieux D. 2000. Winter distribution of woodland caribou in relation to clear-cut logging in west-central Alberta. *Can. J. Zool.* 78(8):1433–40

Smithers LA. 1961. Lodgepole pine in Alberta. Edmonton, Alberta, Canada

Soliveres S, Manning P, Prati D, Gossner MM, Alt F, et al. 2016. Locally rare species influence grassland ecosystem multifunctionality. *Philos. Trans. R. Soc. B Biol. Sci.* 371(1694):1–10

Sorensen T, McLoughlin PD, Hervieux D, Dzus E, Nolan J, et al. 2008. Determining sustainable levels of cumulative effects for boreal caribou. *J. Wildl. Manage.* 72(4):900–905

Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, et al. 2010. Gene movement and genetic association with regional climate gradients in California valley oak (Quercus lobata Née) in the face of climate change. *Mol. Ecol.* 19(17):3806–23

Sparapani R, Spanbauer C, McCulloch R. 2021. Nonparametric machine learning and efficient computation with bayesian additive regression trees: The bart r package. *J. Stat. Softw.* 97(1):1–66

Spear SF, Balkenhol N, Fortin M-J, McRae BH, Scribner K. 2010. Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Mol. Ecol.* 19(17):3576–91

Squires JR, Ruggiero LF. 2007. Winter Prey Selection of Canada Lynx in Northwestern Montana. *J. Wildl. Manage.* 71(2):310–15

St-Louis A, Ouellet J-P, Crête M, Maltais J, Huot J. 2000. Effects of partial cutting in winter on white-tailed deer. *Can. J. For. Res.* 30(4):655–61

Steenweg R, Hebblewhite M, Kays R, Ahumada J, Fisher JT, et al. 2017. Scaling-up camera traps: monitoring the planet’s biodiversity with networks of remote sensors. *Front. Ecol. Environ.* 15(1):26–34

Steinbauer MJ, Dolos K, Reineking B, Beierkuhnlein C. 2012. Current measures for distance decay in similarity of species composition are influenced by study extent and grain size. *Glob. Ecol. Biogeogr.* 21(12):1203–12

Stephens PA, Buskirk SW, Hayward GD, Del Rio CM. 2005. Information theory and hypothesis testing: a call for pluralism. *J. Appl. Ecol.* 42(1):4–12

Stephenson RO, Grangaard D V., Burch J. 1991. Lynx, Felis lynx, predation on red foxes, Vulpes vulpes, caribou, Rangifer tarandus, and Dall sheep, Ovis dalli, in Alaska. *Can. Field-Naturalist*. 105(2):255–62

Stevens VM, Leboulengé É, Wesselingh RA, Baguette M. 2006a. Quantifying functional connectivity: Experimental assessment of boundary permeability for the natterjack toad (Bufo calamita). *Oecologia*. 150(1):161–71

Stevens VM, Polus E, Wesselingh RA, Schtickzelle N, Baguette M. 2004. Quantifying functional connectivity: experimental evidence for patch-specific resistance in the Natterjack toad (Bufo calamita). *Landsc. Ecol.* 19:829–42

Stevens VM, Verkenne C, Vandewoestijne S, Wesselingh RA, Baguette M. 2006b. Gene flow and functional connectivity in the natterjack toad. *Mol. Ecol.* 15(9):2333–44

Stewart FEC, Fisher JT, Burton AC, Volpe JP. 2018. Species occurrence data reflect the magnitude of animal movements better than the proximity of animal space use: *Ecosphere*. 9(2):e02112

Storch I, Lindström E, Jounge J de. 1990. Diet and habitat selection of the pine marten in relation to competition with the red fox. *Acta Theriol. (Warsz).* 35:311–20

Storfer AT, Murphy MA, Evans JS, Goldberg CS, Robinson S, et al. 2007. Putting the “landscape” in landscape genetics. *Heredity.* 98(3):128–42

Storfer AT, Murphy MA, Spear SF, Holderegger R, Waits LP. 2010. Landscape genetics: Where are we now? *Mol. Ecol.* 19(17):3496–3514

Sturbois A, De Cáceres M, Sánchez-Pinillos M, Schaal G, Gauthier O, et al. 2021. Extending community trajectory analysis: New metrics and representation. *Ecol. Modell.* 440:109400

Suárez NM, Betancor E, Fregel R, Rodríguez F, Pestano J, et al. 2012. Genetic signature of a severe forest fire on the endangered Gran Canaria blue chaffinch (Fringilla teydea polatzeki). *Conserv. Genet.* 13(2):499–507

Szabó P, Hédl R. 2011. Advancing the Integration of History and Ecology for Conservation. *Conserv. Biol.* 25(4):680–87

Szucs M, Melbourne BA, Tuff T, Hufbauer RA. 2014. The roles of demography and genetics in the early stages of colonization. *Proc. R. Soc. B Biol. Sci.* 281(1792):20141073

Tallmon DA, Gregovich D, Waples RS, Scott Baker C, Jackson J, et al. 2010. When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Mol. Ecol. Resour.* 10(4):684–92

Tansley AG. 1935. The Use and Abuse of Vegetational Concepts and Terms. *Ecol. Soc. Am.* 16(3):284–307

Tarca AL, Carey VJ, Chen X, Romero R, Drăghici S. 2007. Machine Learning and Its Applications to Biology. *PLoS Comput. Biol.* 3(6):e116

Tattersall ER, Burgar JM, Fisher JT, Burton AC. 2020a. Mammal seismic line use varies with restoration: Applying habitat restoration to species at risk conservation in a working landscape. *Biol. Conserv.* 241:108295

Tattersall ER, Burgar JM, Fisher JT, Burton AC. 2020b. Boreal predator co-occurrences reveal shared use of seismic lines in a working landscape. *Ecol. Evol.* 10(3):1678–91

Taylor PD, Fahrig L, Henein K, Merriam G. 1993. Connectivity is a vital element of landscape structure. *Oikos*. 68(3):571–73

Thomas CD. 2000. Dispersal and extinction in fragmented landscapes. *Proc. Biol. Sci.* 267(1439):139–45

Thompson WL. 2004. *Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters*. Washington DC, USA: Island Press

Thomson AJ, Shrimpton DM. 1984. Weather associated with the start of mountain pine beetle outbreaks. *Can. J. For. Res.* 14:255–58

Tischendorf L, Fahrig L. 2000. On the usage and measurement of landscape connectivity. *Oikos*. 90(1):7–19

Tobler WR. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* 46(June Suppl.):234–40

Toews M, Juanes F, Burton AC. 2017a. Mammal responses to human footprint vary with spatial extent but not with spatial grain. *Ecosphere*. 8(3):e01735

Toews M, Juanes F, Burton AC. 2017b. Mammal responses to human footprint vary with spatial extent but not with spatial grain. *Ecosphere*. 8(3):

Toews M, Juanes F, Burton AC. 2018. Mammal responses to the human footprint vary across species and stressors. *J. Environ. Manage.* 217:690–99

Treml EA, Halpin PN, Urban DL, Pratson LF. 2008. Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landsc. Ecol.* 23(SUPPL. 1):19–36

Trenham PC. 2001. Terrestrial Habitat Use by Adult California Tiger Salamanders. *J. Herpetol.* 35(2):343–46

Trevoy SAL, Janes JK, Sperling FAH. 2018. Where did mountain pine beetle populations in Jasper Park come from? Tracking beetles with genetics. *For. Chron.* 94(1):20–24

Trotter RT, Pepper E, Davis K, Vazquez R. 2019. Anisotropic dispersal by the Asian longhorned beetle (Anoplophora glabripennis): quantifying spatial risk and eradication effort with limited biological data. *Biol. Invasions*. 21(4):1179–95

Tscharntke T, Tylianakis JM, Rand TA, Didham RK, Fahrig L, et al. 2012. Landscape moderation of biodiversity patterns and processes - eight hypotheses. *Biol. Rev.* 87(3):661–85

Turner MG. 1989. Landscape ecology: the effect of pattern on process. *Annu. Rev. Ecol. Syst.* 20:171–97

Turner MG. 2005. Landscape ecology: What is the state of the science? *Annu. Rev. Ecol. Evol. Syst.* 36:319–44

Turner MG, Arthaud GJ, Engstrom RT, Hejl SJ, Liu J, et al. 1995. Usefulness of spatially explicit population models in land management. *Ecol. Appl.* 5(1):12–16

Turner MG, O’Neill R V., Gardner RH, Milne BT. 1989. Effects of changing spatial extent on landscape pattern analysis. *Landsc. Ecol.* 3(3/4):153–62

Tytar V, Nekrasova O, Pupins M, Skute A, Marushchak O, et al. 2021. Identifying Environmental Refuges (" Coldspots ") from Infection by Batrachochytrium Dendrobatidis of Amphibians in Eastern Europe. *1st Int. Electron. Conf. Biol. Divers. Ecol. Evol.*

van Strien MJ. 2017. Consequences of population topology for studying gene flow using link-based landscape genetic methods. *Ecol. Evol.* 7(14):5070–81

Van Strien MJ, Keller D, Holderegger R. 2012. A new analytical approach to landscape genetic modelling: Least-cost transect analysis and linear mixed models. *Mol. Ecol.* 21(16):4010–23

Veech JA. 2012. Significance testing in ecological null models. *Theor. Ecol.* 5(4):611–16

Velázquez E, Martínez I, Getzin S, Moloney KA, Wiegand T. 2016. An evaluation of the state of spatial point pattern analysis in ecology. *Ecography.* 39(11):1042–55

Vellend M, Baeten L, Becker-Scarpitta A, Boucher-Lalonde V, McCune JL, et al. 2017. Plant Biodiversity Change Across Scales During the Anthropocene. *Annu. Rev. Plant Biol.* 68:563–86

Verhulst B. 2014. In defense of P values. *Ecology*. 95(3):611–17

Vincent H, Bornand CN, Kempel A, Fischer M. 2020. Rare species perform worse than widespread species under changed climate. *Biol. Conserv.* 246:108586

Wagner HH, Dray S. 2015. Generating spatially constrained null models for irregularly spaced data using Moran spectral randomization methods. *Methods Ecol. Evol.* 6(10):1169–78

Wagner HH, Fortin M-J. 2013. A conceptual framework for the spatial analysis of landscape genetic data. *Conserv. Genet.* 14(2):253–61

Waller NL, Gynther IC, Freeman AB, Lavery TH, Leung LKP. 2017. The Bramble Cay melomys Melomys rubicola (Rodentia:Muridae): A first mammalian extinction caused by human-induced climate change? *Wildl. Res.* 44(1):9–21

Walpole AA, Bowman J, Murray DL, Wilson PJ. 2012. Functional connectivity of lynx at their southern range periphery in Ontario, Canada. *Landsc. Ecol.* 27(5):761–73

Wang IJ, Bradburd GS. 2014. Isolation by Environment. *Mol. Ecol.* n/a-n/a

Wang IJ, Savage WK, Bradley Shaffer H. 2009. Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (Ambystoma californiense). *Mol. Ecol.* 18(7):1365–74

Wang IJ, Shaffer HB. 2017. Population genetic and field-ecological analyses return similar estimates of dispersal over space and time in an endangered amphibian. *Evol. Appl.* 10(6):630–39

Wang X, Tejedor A, Wang Y, Moreno Y. 2021. Unique superdiffusion induced by directionality in multiplex networks. *New J. Phys.* 23(1):013016

Wang Y, Hogg EH, Price DT, Edwards J, Williamson T. 2014. Past and projected future changes in moisture conditions in the Canadian boreal forest. *For. Chron.* 90(5):678–91

Wang Y, Naumann U, Wright ST, Warton DI. 2012. Mvabund- an R package for model-based analysis of multivariate abundance data. *Methods Ecol. Evol.* 3(3):471–74

Waples RS. 2005. Genetic estimates of contemporary effective population size: To what time periods do the estimates apply? *Mol. Ecol.* 14(11):3335–52

Warton DI, Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, et al. 2015. So Many Variables: Joint Modeling in Community Ecology. *Trends Ecol. Evol.* 30(12):766–79

Warton DI, Lyons M, Stoklosa J, Ives AR. 2016. Three points to consider when choosing a LM or GLM test for count data. *Methods Ecol. Evol.* 7(8):882–90

Warton DI, Wright ST, Wang Y. 2012. Distance-based multivariate analyses confound location and dispersion effects. *Methods Ecol. Evol.* 3(1):89–101

Wasser SK, Keim JL, Taper ML, Lele SR. 2011. The influences of wolf predation, habitat loss, and human activity on caribou and moose in the Alberta oil sands. *Front. Ecol. Environ.* 9(10):546–51

Watts PC, Rousset F, Saccheri IJ, Leblois R, Kemp SJ, Thompson DJ. 2007. Compatible genetic and ecological estimates of dispersal rates in insect (Coenagrion mercuriale: Odonata: Zygoptera) populations: Analysis of "neighbourhood size" using a more precise estimator. *Mol. Ecol.* 16(4):737–51

Wearn OR, Carbone C, Rowcliffe JM, Pfeifer M, Bernard H, Ewers RM. 2019. Land-use change alters the mechanisms assembling rainforest mammal communities in Borneo. *J. Anim. Ecol.* 88(1):125–37

Weber B. 2017. Pine beetles from Jasper National Park now infesting commercial forest. *Can. Broadcast. Corp.*, . https://www.cbc.ca/news/canada/edmonton/pine-beetles-from-jasper-national-park-moving-into-commercial-forest-1.4406820

Webster MS, Marra PP, Haig SM, Bensch S, Holmes RT. 2002. Links between worlds: Unraveling migratory connectivity

Weir BS, Cockerham CCC. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution.* 38(6):1358–70

Welch H, Brodie S, Jacox MG, Bograd SJ, Hazen EL. 2020. Decision-support tools for dynamic management. *Conserv. Biol.* 34(3):589–99

Wenger SJ, Olden JD. 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol. Evol.* 3(2):260–67

Westfall PH, Young SS, Lin DKJ. 1998. Forward selection error control in the analysis of supersaturated designs. *Stat. Sin.* 8(1):101–17

White JW, Rassweiler A, Samhouri JF, Stier AC, White C. 2014. Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos*. 123(4):385–88

Whitlock MC, McCauley DE. 1999. Indirect measures of gene flow and migration: FST not equal to 1/(4Nm + 1). *Heredity (Edinb).* 82(November 1998):117–25

Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP. 2006. Why do we still use stepwise modelling in ecology and behaviour? *J. Anim. Ecol.* 75(5):1182–89

Whittington J, Hebblewhite M, Decesare NJ, Neufeld L, Bradley M, et al. 2011. Caribou encounters with wolves increase near roads and trails: a time-to-event approach. *J. Appl. Ecol.* 48(6):1535–42

Wiens JA. 1989. Spatial Scaling in Ecology. *Funct. Ecol.* 3(4):385–97

Wiens JA. 1992. What is landscape ecology, really? *Landsc. Ecol.* 7(3):149–50

Wilkinson DP, Golding N, Guillera-Arroita G, Tingley R, McCarthy MA. 2019. A comparison of joint species distribution models for presence–absence data. *Methods Ecol. Evol.* 10(2):198–211

Wilkinson L. 1979. Tests of significance in stepwise regression. *Psychol. Bull.* 86(1):168–74

Wilkinson L, Dallal GE. 1981. Tests of significance in forward selection regression with an f-to-enter stopping rule. *Technometrics*. 23(4):377–80

Willi Y, Griffin P, Van Buskirk J. 2013. Drift load in populations of small size and low density. *Heredity.* 110(3):296–302

Winder R, Stewart FEC, Nebel S, McIntire EJB, Dyk A, Omendja K. 2020. Cumulative Effects and Boreal Woodland Caribou: How Bow-Tie Risk Analysis Addresses a Critical Issue in Canada’s Forested Landscapes. *Front. Ecol. Evol.* 8:1

Winegardner AK, Legendre P, Beisner BE, Gregory-Eaves I. 2017. Diatom diversity patterns over the past c. 150 years across the conterminous United States of America: Identifying mechanisms behind beta diversity. *Glob. Ecol. Biogeogr.* 26(11):1303–15

Winiarski KJ, Peterman WE, McGarigal K. 2020. Evaluation of the R package 'resistancega': A promising approach towards the accurate optimization of landscape resistance surfaces. *Mol. Ecol. Resour.* 20(6):1583–96

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol. Rev.* 88(1):15–30

With KA, Gardner RH, Turner MG. 1997. Landscape Connectivity and Population Distributions in Heterogeneous Environments. *Oikos*. 78(1):151

Wittische J, Heckbert S, James PMA, Burton AC, Fisher JT. 2021. Community-level modelling of boreal forest mammal distribution in an oil sands landscape. *Sci. Total Environ.* 755(2):142500

Wittische J, Janes JK, James PMA. 2019. Modelling landscape genetic connectivity of the mountain pine beetle in western Canada. *Can. J. For. Res.* 49(11):1339–48

Wood C, Qiao Y, Li P, Ding P, Lu B, Xi Y. 2010. Implications of rice agriculture for wild birds in China. *Waterbirds*. 33(SPEC.ISSUE.1):30–43

Wood SLR, Martins KT, Dumais-Lalonde V, Tanguy O, Maure F, et al. 2021. Missing interactions: the current state of multispecies connectivity analysis. *bioRxiv Ecol.*

Wright S. 1943. Isolation by Distance. *Genetics*. 28(2):114–38

Wu J. 2004. Effects of changing scale on landscape pattern analysis: Scaling relations. *Landsc. Ecol.* 19(2):125–38

Yanai RD, See CR, Campbell JL. 2018. Current Practices in Reporting Uncertainty in Ecosystem Ecology. *Ecosystems*. 21(5):971–81

Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, et al. 2018. Outstanding hallenges in the transferability of ecological models. *Trends Ecol. Evol.* 33(10):790–802

Yeaman S. 2015. Local adaptation by alleles of small effect. *Am. Nat.* 186(Oct. Suppl. 1):S74–89

Yee TW. 2015. *Vector Generalized Linear and Additive Models: With an Implementation in R*. New York, New York, USA: Springer International Publishing AG

Yee TW. 2020. The VGAM package for negative binomial regression. *Aust. New Zeal. J. Stat.* 62(1):116–31

Yee TW. 2021. *Complements to Vector Generalized Linear and Additive Models*. New York, New York, USA: Springer International Publishing AG

Yeung ACY, Paltsev A, Daigle A, Duinker PN, Creed IF. 2019. Atmospheric change as a driver of change in the Canadian boreal zone. *Environ. Rev.* 27(3):346–76

Yu H, Nason JD, Ge X, Zeng J. 2010. Slatkin's Paradox: When direct observation and realized gene flow disagree. A case study in Ficus. *Mol. Ecol.* 19(20):4441–53

Zeller KA, Jennings MK, Vickers TW, Ernest HB, Cushman SA, Boyce WM. 2018. Are all data types and connectivity models created equal? Validating common connectivity approaches with dispersal data. *Divers. Distrib.* 24(7):868–79

Zeller KA, McGarigal K, Whiteley AR. 2012. Estimating landscape resistance to movement: A review. *Landsc. Ecol.* 27(6):777–97

Zeller KA, Wattles DW, Destefano S. 2020. Evaluating methods for identifying large mammal road crossing locations: black bears as a case study. *Landsc. Ecol.* 35(8):1799–1808

Zipkin EF, Andrew Royle J, Dawson DK, Bates S. 2010. Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biol. Conserv.* 143(2):479–84