JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE

234

Anne Laita

Conservation in space

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 234

Anne Laita

Conservation in Space

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa S212
tammikuun 20. päivänä 2012 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in building Seminarium, auditorium S212, on January 20, 2012 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2012

Conservation in Space

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 234

Anne Laita

Conservation in Space

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2012

Editors
Jari Haimi
Department of Biological and Environmental Science, University of Jyväskylä
Pekka Olsbo, Harri Hirvi
Publishing Unit, University Library of Jyväskylä

Cover picture: Air perspective on the landscape mosaic surrounding my grandmother´s place

URN:ISBN:978-951-39-4618-0
ISBN 978-951-39-4618-0 (PDF)

ISBN 978-951-39-4617-3 (nid.)
ISSN 1456-9701

Jyväskylä University Printing House, Jyväskylä 2012

”Hiiri mittaa maailmaa männynneulasella”
-Hannele Huovi

ABSTRACT

Laita, Anne
Conservation in space
Jyväskylä: University of Jyväskylä, 2012, 41 p.
(Jyväskylä Studies in Biological and Environmental Science
ISSN 1456-9701; 234)
ISBN 978-951-39-4617-3 (nid.)
ISBN 978-951-39-4618-0 (PDF)
Yhteenveto: Lajien suojelu tilassa
Diss.

Species distributions are highly spatially structured owing to three primary reasons. First, the spatial structure in environmental factors is passed to species distributions. Second, species respond to their environments in a hierarchical manner – at the spatial scales rising from species ecology. Third, ecological processes themselves are affected by space as the interactions between organisms are, to a high degree, limited by distance. Explicit consideration of space as an important underlying force affecting ecological processes and patterns does not only provide new theoretical and empirical knowledge but offers tools for conservation planning that targets for the long-term persistence of biodiversity. In this thesis, I have examined two current conservation issues from a spatial perspective and at the landscape level. First, I studied the role of woodland key habitats as an additional part of the traditional protected area network. My results showed that dispersal ability of species is linked with the potential conservation merits of woodland key habitats. The protected habitat area served as a misleading measure for the amount of protected habitat actually available to species at the landscape level. Second, I studied the habitat associations of the Siberian Jay (*Perisoreus infaustus*) in boreal forest landscapes by using spatially explicit species distribution modeling (SDM) methods with a multi-scale approach. My results showed that the species distribution is affected by the landscape context, and is most probably shaped by the intrinsic aggregation rising from species behavior. The comparisons of various spatial methods indicated that their conceptual underpinnings are passed on to the results they are producing, which emphasizes the role of understanding theoretical assumptions and the intricacies of spatial effects when interpreting spatial results. My results add to the accumulated scientific evidence for the notable role of space affecting species living in fragmented landscapes, and emphasize the importance of spatial considerations in conservation planning.

Keywords: Connectivity; graph theory; hierarchical model; Siberian Jay; spatial autocorrelation; species distribution model; woodland key habitat.

Author's address Anne Laita
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland
anne.laita@jyu.fi

Supervisors Professor Mikko Mönkkönen
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Professor Janne S. Kotiaho
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Reviewers PhD Lluís Brotons
Centre Tecnològic Forestal de Catalunya
Pujada del Seminari s/n
25280 Solsona
Spain

Professor Miska Luoto
Department of Geosciences and Geography
P.O.Box 64
FI-00014 University of Helsinki
Finland

Opponent Professor Otso Ovaskainen
Department of Biological and Environmental Sciences
P.O. Box 65
FI-00014 University of Helsinki
Finland

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original papers, which will be referred to in the text by their Roman numerals I-IV.

I Laita, A., Mönkkönen, M. & Kotiaho, J.S. 2010. Woodland key habitats evaluated as part of a functional reserve network. Biological Conservation 143: 1212-1227.

II Laita, A., Kotiaho, J.S. & Mönkkönen, M. 2011. Graph-theoretic connectivity measures: what do they tell us about connectivity? Landscape Ecology 26: 951-967.

III Laita, A. & Mönkkönen, M. 2011. Space, scales and imperfect detection: modeling the habitat association of the Siberian Jay (*Perisoreus infaustus)* in boreal forest landscapes. Submitted manuscript.

IV Laita, A. & Mönkkönen, M. 2011. Accounting for space in species distribution modeling – alternative routes leading to a same destination? Manuscript.

The table shows the contributions to the original papers. Smaller contributions are stated in the acknowledgements of the original papers.

	I	II	III	IV
Original idea	MM, AL	AL	MM, AL	AL
Analyses	AL	AL	AL	AL
Writing	AL, MM, JSK	AL, MM, JSK	AL, MM	AL, MM

AL = Anne Laita, MM = Mikko Mönkkönen, JSK = Janne S. Kotiaho

TERMINOLOGY

The glossary briefly describes the terms that are used in the thesis. For the terms, essential references are also provided.

Connectivity	Potential for dispersal; rises in interaction between the landscape structure and species ecological traits (Taylor et al. 1993).
Fragmentation	Dissection of habitat into smaller and more distant parcels than postulated basis on the habitat loss alone (Fahrig 1997, Fahrig 2003).
Habitat loss	Loss of habitat for a particular species; comes with many changes in landscape structure and configuration, e.g., larger inter-patch distances and smaller patch sizes (Fischer & Lindenmayer 2007).
Hierarchical scales	Species-specific assemblages of spatial scales at which the species respond to environment, the operation scales of different mechanisms underlying species distributions (Wiens 1989, Levin 1992).
Matrix	Landscape area that surrounds habitat patches; non-habitat (see Haila 2002).
Matrix effects	General term for the great variety of ways and mechanisms that the matrix exerts on both ecological quality of individual habitat patches as well as interpatch movements (see Bender & Fahrig 2005, Fischer & Lindenmayer 2007).
Spatial autocorrelation	Correlation of a response variable with itself in space; correlation may be produced by extrinsic (thereby rising correlation may be also referred as spatial dependence) and intrinsic factors rising from species ecology (Legendre 1993).
Species distribution model (SDM)	"A model that relates species distribution data (occurrence or abundance at known locations) with information on the environmental and/or spatial characteristics of those locations", as defined by Elith & Leathwick (2009) (see also Guisan & Zimmermann 2000, Austin 2002, Guisan & Thuiller 2005).

1 INTRODUCTION

1.1 Conservation goals

Habitat loss, degradation and fragmentation are the key drivers of biodiversity loss at scales ranging from global to local ones (e.g., Fischer and Lindenmayer 2007). Fragmentation means dissection of habitat to smaller and more isolated patches, and although it operates usually in concert with habitat loss, it has also independent adverse effect on species (e.g, Villard et al. 1999, Wiegand et al. 2005, but see also Trzcinski et al. 1999). There are theoretical and empirical evidence that the effect of fragmentation (independently from the effect of habitat loss) intensifies when the amount of habitat falls below a critical species- and landscape-specific threshold (Andrén 1994, Mönkkönen and Reunanen 1999, With and King 1999, Fahrig 2001). As the amount of suitable habitat in a landscape is globally very low for many specialized species, it turns essential to take actions against fragmentation as part of conservation planning. Conservation that successfully and in an ecologically founded way operates on these three components – habitat amount, quality and the spatial configuration of habitat at the landscape level – forms the essential basis for halting the loss of biodiversity and safeguarding the persistence of biodiversity in the long run.

Successful biodiversity conservation poses two requirements for a reserve network (sensu Margules & Pressey 2000). First is representativeness, a widely acknowledged goal of encompassing a full variety of biodiversity. Second, much less considered goal, is species persistence, i.e., a reserve network should maintain its biodiversity values also in the long run. Many pursuits of evaluating or supplementing existing networks have concentrated on the criteria of representativeness. There exist many kinds of site selection algorithms that aim at maximizing the species pool residing in the protected area network based on different criteria. However, the criteria of representativeness and persistence may be in conflict; the realization of the criteria of representativeness may result in overly small population sizes and non-functional, i.e. weakly connected, protected area networks that do not assure species persistence in a best possible way. The goal of ensuring species persistence is implicitly or explicitly underlying various spatial considerations in the designation of a conservation area network (Araújo & Williams 2000). The integration of spatial criteria to site-selection algorithms may be used to emphasize the objective of population persistence (e.g., Cabeza 2003, Cabeza and Moilanen 2003, Williams et al. 2005, Önal & Briers 2005, van Teeffelen et al. 2006, Rayfield et al. 2009). So, instead of safeguarding the realized pattern, the spatial considerations in conservation shift the focus more to arranging preconditions for ecological processes to operate. Usually, considering the effect of space on ecological processes leads to more aggregated conservation networks and acts than carried out on pure non-spatial grounds.

Conservation of ecological processes necessitates the spatial and temporal domains of conservation to be extended (Angelstam et al. 2004). Spatial considerations are essentially linked with the temporal dimension of conservation, because habitat provided with suitable configuration implicitly provides potential for the ecological processes to operate over time. Landscape scale exhibits the cumulative effect of local level activities, and is the scale over which activities affect the persistence of populations. Landscape-scale represents an increasingly more important scale of conservation management, because the whole cannot be understood based on distinct elements but as an interrelated system. This calls for the integrated view on landscapes and their management.

1.2 Theoretical basis of spatial ecology

Many important ecological theories have a spatial dimension in their cores, including e.g., theory of island biogeography (MacArthur and Wilson 1967), the source-sink theory (Pulliam 1988) and metapopulation theory (Hanski 1999). This underlines the effect of space in shaping ecological patterns and processes. Conservation planning and reserve design has been influenced and shaped by these theories for many decades (see e.g., Diamond 1975). These theories provide conceptual answers for the allocation of conservation efforts in space.

Metapopulation theory, focusing on the dynamics of spatially interrelating populations, is the leading theory of spatially oriented ecology. Theoretical and empirical work rooted in metapopulation theory has shown that the spatial structure of ecological interactions has a profound effect on population dynamics (Hanski 1998). Dispersal between habitat patches has major influence on the rates of patch colonization and extinction, and is thus essentially linked with population dynamics in fragmented landscapes. Combining this spatially driven component with the models on within-patch dynamics allows to end up with the evaluation for the long-term persistence of a species in a landscape.

The theoretical background for spatial ecology offers mainly a dichotomous view on landscapes – habitat patches interspersed in the matrix formed by non-habitat. Within this spatial theoretical framework, the sizes and configuration of habitat patches are the main drivers of spatial ecological effects. However, patch size and isolation have not always proved to be especially powerful predictors of species occupancy, which has been deduced to result from the strong ecological effect of the matrix that surrounds habitat patches (see e.g., Bender & Fahrig 2005, Prugh et al. 2008, Watling et al. 2011). Depending on the species ecology, the habitat/non-habitat dichotomy may be an oversimplification of the landscape as actually experienced by organisms.

The field of landscape ecology emphasizes the interaction between spatial pattern and ecological process across a range of scales (Turner at al. 2001). Although landscape ecology does not provide any well-defined and concrete theoretical theoretical framework, it is the discipline that has greatly increased the knowledge on species living in heterogeneous and spatially structured landscapes. Landscape ecological research has provided insights into the complexities of real landscapes, and profound influences of these landscape structures on species. Landscape ecology as a discipline has also raised the concept of scale to one of the central tenets of ecological phenomena (Levin 1992, Schneider 2001, Dungan et al. 2002).

1.3 Species embedded in landscapes

Matrix surrounding habitat patches may have various effects on ecological responses, and those effects may be mediated by several mechanisms. All these effects are missed if landscapes are viewed only within a dichotomous framework. Habitat patches are not in many cases embedded in a totally inhospitable matrix but are surrounded by landscape elements providing different kinds of supplemental resources as well as varying permeability for species dispersal (see Ricketts 2001, Manning et al. 2004). In addition to shaping effective isolation between habitat patches, matrix may also have additional profound influences on the species living in a habitat fragments. For example, matrix may cast effects on habitat patches through edge effects (Harrison & Bruna 1999), or by altering biotic interactions (Cronin 2007). Matrix has been shown to influence biodiversity values of habitat patches in a wealth of studies (for a review, see Prevedello & Vieira 2010). Appreciating the importance of matrix affecting ecological processes, conservation through distinct parcels of protected habitat should be increasingly replaced by the landscape-scale conservation strategies that would safeguard ecological variability in space and time (Lindenmayer & Franklin 2002, Villard & Jonsson 2009).

Species respond to their environments in a hierarchical manner and over assemblages of species-specific spatial scales (Wiens 1989, Levin 1992), which further inflates the role of the larger spatial context in shaping biodiversity values at local levels. Many multi-scale studies have shown that the occurrence and abundance of species cannot be explained based on local factors alone, but also broader scales are important (Drapeau et al. 2000, Bakker et al. 2002, Cushman & McGarigal 2004). The relative importance of different environmental variables on determining species distributions changes with the spatial scale of analysis (see e.g., Buler et al. 2007). Ecological processes may show differences in their responses to hierarchical scales. For example, Sirami et al. (2008) showed that colonization events in a bird community were multi-scale dependent while the extinction was driven solely by local habitat factors.

The response of the species to environmental factors beyond the local ones has important implications for the scale of species conservation. In many cases, there might be a mismatch between the scale over which species respond to their environments and the scale over which conservation is planned (e.g., Nilsson 2009, Villard & Jonsson 2009). Usually this means that the conservation is planned at the scale notably smaller than that of species environmental response. This can rise because of the practical conservation constraints set by the landownership structure. For example, high share of private ownerships in a landscape may force conservation planning to a spatial scale unavoidably too small with respect to domain of ecological responses. However, the mismatch of scales may also be due to the lack of knowledge and expertise to view the ecological processes so tightly and in a rich manner operating in space.

1.4 Species distributions exhibit aggregation

Ecological responses (e.g. species occurrence, abundance, species richness) usually exhibit positive spatial autocorrelation so that nearby locations have more similar response values than those further apart. This similarity may rise mainly via two different mechanisms - extrinsic and intrinsic factors (Legendre 1993, Fortin & Dale 2005). Extrinsic factors (such as climate, land cover and forest characteristics) that either indirectly or directly control ecological processes are usually themselves spatially structured. The positive autocorrelation in their structure is then transferred to the ecological response. This extrinsic form of autocorrelation may be referred to as *spatial dependence* to distinguish it from the true autocorrelation (correlation of a response variable with itself) rising from intrinsic factors (Legendre 1993, Lichstein et al. 2002) The intrinsic factors leading to spatial autocorrelation in a response (over the control of environmental factors) are generated by species ecology. The most evident intrinsic factor underlying spatial autocorrelation in species distributions is dispersal that is limited by distance (e.g., Bahn 2008). In addition, many kinds of ecological factors like biotic interactions (competition and predation) and social behavior may generate aggregation in a response (for a list of extrinsic and intrinsic factors producing positive autocorrelation in a response, see Miller et al. 2007).

The spatial dependence and autocorrelation in a response is usually seen as a statistical nuisance, because spatial autocorrelation poses many kinds of problems for statistical testing, parameter estimation and model selection. The statistical problems are caused by the spatial dependence and autocorrelation that is left to model residuals after environmental covariates are included to a model. The residual autocorrelation may rise as an outcome of several intrinsic and extrinsic factors, e.g., missing environmental covariates (Ishihama et al. 2010), misidentified shape of response (Santika & Hutchinson 2009) and observation bias. Spatial autocorrelation represents one form of pseudoreplication so that samples provide partly overlapping information, and an individual sample does not contribute information worth of full degree of freedom (e.g, Dale & Fortin 2002, Fortin & Dale 2005). Because of this sample dependence, the degrees of freedom are overrated in non-spatial statistical analyses. This leads to too small standard error for the parameters, and thus inflation of Type I error rate (Dormann et al. 2007, Hoeting 2009). Non-spatial statistical tests carried out in the presence of positive spatial autocorrelation thus lead to too liberal test results (Cliff & Ord 1981). Spatial autocorrelation may also cause shifts in coefficient values (Dormann 2007, Bini et al. 2009), even inverting the direction of relationship (Kühn 2007). The relative importance of explanatory variables has also been shown to be affected when switched from non-spatial models to spatial ones, which have implications for the model selection in the presence of autocorrelation (Tognelli & Kelt 2004, Hoeting et al. 2006, Segurado et al. 2006, Diniz-Filho et al. 2008). As results from statistical testing and modeling essentially guide conservation practices, the problems of inference in a presence of autocorrelation are not trivial. Overall, accounting for the spatial autocorrelation allows for accurate and valid inferences made about the species-environment relationships (Keitt et al. 2002), and provides also sound basis for conservation management.

The rigorous control of spatial autocorrelation has, however, merits far beyond the better meeting of statistical assumptions. As different spatial modeling methods have become widely available, with accumulated experiences and knowledge of their use, the understanding of the nuanced effects of space has risen to a new level. The capture of spatial autocorrelation leads usually to better predictive performance of the distribution models compared to non-spatial ones (Augustin et al. 1996, Knapp et al. 2003, Wintle & Bardos 2006, Hoeting 2009). This indicates that, in many cases, species distributions cannot be explained and predicted without explicit consideration of spatial effects. There may lie important information in spatial autocorrelation, the explicit use of which may offer qualitative and quantitative insights to the aggregating forces underlying patterns. The range and strength of spatial autocorrelation may serve as important cues to the nature and scale of process creating the patterns. It has also been suggested that accounting for the spatial autocorrelation may allow unravelling the hierarchical nature of environmental control, and identifying the environmental factors that act at finer spatial scales (Diniz-Filho et al. 2003).

Further, comparisons of results from different spatial methods, or from the variations within one method, can be used to gain information on the role of intrinsic and extrinsic factors leading to aggregation (van Teeffelen & Ovaskainen 2007). However, the full quantification of their independent effects on aggregation is not possible. The interpretations in favour of intrinsic autocorrelation are strengthened by the identification of plausible, species-specific ecological traits and mechanisms expected to produce aggregation. The identification of intrinsic spatial factors behind aggregation would be especially important from the point of view of species conservation, because this offers a founded basis for the level and scale of aggregation to be targeted with spatial conservation planning. The aggregation produced by intrinsic factors is the element that cannot be understood and taken into account in conservation leaning on environmental factors alone, but it really creates needs for rigorous spatial thinking and planning.

1.5 Connectivity

Connectivity, in short, describes the potential for the intrinsic factors, which are mediated by dispersal, to operate. Ultimately, connectivity can be viewed as a realization of dispersal events (for example, as verified by population genetic analyses, see Segelbacher et al. 2010). Habitat connectivity is a prerequisite for the operation of ecological processes in space and time. Connectivity has been empirically shown to have positive effect on patch occupancy (e.g., Brouwers & Newton 2009b, Ranius et al. 2010), patch colonization (Hanski 1999, van Langevelde 2000), population size and density (Hanski et al. 1994), population growth rate, genetic variability, species persistence and species richness (Brückmann et al. 2010), and negative effect on patch extinction (Franken & Hik 2004). Connectivity can be viewed as a key concept in spatially oriented ecology, and it serves an essential basis for spatial conservation planning.

Connectivity is an emergent property, rising in an interaction between the physical structure of a given landscape and the ecological properties of a species (i.e. dispersal ability and propensity, habitat specialization, social system) so that landscape connectivity is always very species-specific phenomenon, not describable by a single value (Merriam 1984, Taylor et al. 1993, Crooks & Sanjayan 2006). For the part of the physical landscape structure, connectivity embraces two components, habitat amount (and quality) and its configuration, while still allowing explicit considerations of both. Connectivity is not, however, only a function of habitat amount and configuration, but encompasses ultimately a complexity of other factors of how organisms interact with spatial heterogeneity in a landscape (e.g., properties of a matrix) (Gustafsson & Gardner 1996, Mönkkönen & Reunanen 1999). Depending on species ecological traits, composition and configuration of the landscape mosaic surrounding habitat patches may have notable effects on species perception of connectivity. In these cases, straight-line distances between habitat patches may serve as a misleading and overly optimistic measure of the inter-patch separation actually experienced by a species (e.g., Chardon et al. 2003).

The concept of connectivity has been approached from several (partly conflicting) perspectives during the last decade (see Tischendorf & Fahrig 2000, Moilanen & Hanski 2001, Tischendorf & Fahrig 2001), much reflecting the complexity of the concept. Indeed, deep down the concept integrates much of the accumulated theoretical and empirical knowledge on species living in spatially heterogeneous, fragmented landscapes. The species specific perception of connectivity turns also the evaluation and management of landscape connectivity into challenging task. Differences in dispersal modes between group of organisms (i.e. whether they are active vs. passive dispersers, have knowledge of their surroundings) may affect the suitability of different connectivity measures (e.g., Ranius et al. 2010). From the management point of view, the optimal connectivity conservation will almost inevitably differ between species. For many species, the connectivity conservation does not only include habitat conservation and spatial planning, but also the management efforts to improve the quality of matrix.

Species dispersal and conservation

Species dispersal has been attributed to have a major role in shaping species distributions (Bahn 2008) and affecting population dynamics (Hanski 1994, Hanski 1999) in fragmented landscapes. Species dispersal ability is an important ecological trait in affecting connectivity responses of a species. Dispersal ability of a species may thus have implications for the optimal conservation strategy applied. However, there still exists a poor knowledge of the quantitative aspects of dispersal (range and rate) for most species (but see e.g., Edman et al. 2004b, Öckenger et al. 2005, Ranius 2006, Brouwers & Newton 2009a). Dispersal ability of a species is not even quantifiable in a straightforward way because it is a realization of many kinds of ecological and behavioral traits. The dispersal ability of passively dispersing species is difficult to determine. For example, Schleicher et al. (2011) found out in their study on plants living in urban fragments that positive connectivity response was not only linked with dispersal range, but also with the number of seeds produced by an individual plant. On the other hand, for actively dispersing species the properties of matrix and behavioral traits linked with dispersal may play an important role in their connectivity responses.

The role of stochastic dispersal events (e.g. extreme wind conditions) in species dispersal is dependent on temporal and spatial scale. At the very local scale and in a narrow timeframe, the probability of extreme dispersal events is low. Rare events, however, turn more probable when the spatial and temporal frames are broadened. From the point of view of species conservation, the emphasis should be on the more systematic component of species dispersal rather than the more unpredictable one, driven by stochastic extremes. In wind-dispersed wood-decaying fungi with a high potential dispersal range, local spore sources have been still found to drive colonization patterns (Edman et al. 2004a).

Species with poor dispersal ability are suggested to be especially sensitive to habitat fragmentation (e.g., Henle et al. 2004). Range of dispersal sets the species-specific scale at which the effects of fragmentation are experienced. Susceptibility to isolation has been shown to be associated with the conservation status of species (Penttilä et al. 2006). However, dispersal ability of a species is still one ecological characteristics among the others, and for example, competitive ability of a species and intrinsic rate of population growth may have as important role as dispersal on the success of populations in fragmented landscapes. The relative role of these different ecological traits on the population dynamics of species will probably vary with respect to landscape context.

1.6 Preparing for the future

Conservation planning is at its best proactive, preparing for the expected and unexpected environmental changes brought about by the future. Spatial considerations in conservation may play a great role in our successful preparation for these changes. Global climate change is predicted to pose a major threat to global biodiversity over the next century. Landscape fragmentation may further magnify the effects of climate change, because the range shifts may be restricted in highly fragmented landscapes. This interaction between climate and habitat fragmentation is empirically supported, e.g., by Hill et al. (1999), Hill et al. (2001) and Melles et al. (2011) (see also Opdam & Wascher 2004), and may be especially profound for species with low dispersal ability (Pearson and Dawson 2005). The predictions for range shifts based purely on climatic variables may thus be overly optimistic. Landscape connectivity will not only be related to species' abilities to track shifts in their habitats, but also their abilities to adapt to new conditions and their persistence in the new landscapes they will colonize (Taylor 2006) Connectivity management may thus be justifiably seen as one important mean to mitigate some of the negative biodiversity effects caused by the climate change.

Designation of habitat networks that would be resistant to environmental changes and population extirpations poses great challenges to spatial conservation planning. Resistant networks possess redundant connections, so that passages between habitat patches are not dependent on one route only but there exist alternative connections. This is a topic largely ignored so far, although the importance of habitat network robustness is already acknowledged (Williams et al. 2005, Matisziw & Murray 2009, Spring et al. 2010) Long-term management of connectivity is a more complex practice than managing connectivity in one point of time (Matisziw & Murray 2009).

2 AIMS OF THE THESIS

My thesis aims at understanding the multiple effects of space behind species distributions and their implications for species conservation. My thesis consists of four papers (I-IV), all covering space considerations and associated methodology from different angles. I have viewed and evaluated conservation practices at the landscape level and in a spatial context. When comparing different spatial methods (i.e., connectivity measures and spatial modelling methods), my main interest has been in the better understanding of a phenomenon itself, not the mechanistic comparisons of the measures as such. I have viewed the space effects through connectivity concept as well as trough spatial autocorrelation and hierarchical scale effects, finally bringing them altogether in the last study. In the first two papers (I –II), I had a look on a forested landscape from a binary perspective – protected reserve areas embedded in unprotected matrix. Whereas in the consequent papers (IIII –IV), I have viewed the landscapes as spatially heterogeneous continuities, composed of different cover types. Here, I undergo the general background and main questions behind each study.

In the first study (I), I addressed contributions of so called woodland key habitats (WKHs) to landscape level connectivity. WKHs are small-sized habitat patches whose natural characteristics the Finnish Forest Act obliges to be maintained. WKHs are embedded in managed forest landscapes, and represent a branch of new forestry practices that aim at integrating biodiversity concerns to economically profitable forest management (Hansson 2001, Timonen et al. 2010). At the landscape level, they form a protected area network with traditional reserves which represent the primary mean of biodiversity protection. In this study, my interests focused on evaluating the protected area network from a spatial perspective, and to understand the connectivity merits and spatial roles of the WKH patches as additional participants in that network.

In the second study (II), I used the networks of the previous study (I) as a testing platform for the performance of various connectivity measures rooted in graph theory (see the Materials and Methods part). An increasing interest in landscape connectivity has been accompanied with a proliferation of various connectivity measures, rising from different theoretical backgrounds and uses (Kindlmann & Burel 2008). I wanted to investigate whether the different connectivity measures value the connectivity contributions of the WKHs consistently at the level of the whole landscape and individual patches. I also wanted to test the role of network density (i.e. the amount of patches per area) as one potential factor leading to possible evaluational disparity among measures.

In the third study (III), I pursued to shed light on the habitat association of the Siberian Jay (*Perisoreus infaustus)* at the territory and landscape level, because these broader spatial scales of association are rather poorly known for this species. In southern Finland, the species occurs at the southern range of its distribution area. The declines of the species during recent decades have been particularly directed to the southern populations, whose current distribution pattern has turned to highly patchy (Väisänen et al. 1998). The declines have been attributed to commercial forestry, and the conservation of the species is a hot and current issue in Finland. Social behavior of the species is complex and its dispersal is restricted to rather short distances. Therefore, I wanted to examine the additional power of spatial methods at explaining the recent distribution pattern in the southern Finland.

In the fourth study (IV), I compared the parameter estimation and predictive performance of two spatial modeling methods with Bayesian implementation. Modeling results for the binary response variable, like species occurrence, has been found to be especially sensitive to the choice of spatial modeling method (Dormann et al. 2007). For that reason, I wanted to investigate whether this also holds for the two different spatial approaches fitted on the Siberian Jay data of the previous study (III). I was also interested in pondering the reasons behind the possible differences in the parameter estimates and predictions. As connectivity represents potential for intrinsic processes (mediated by dispersal) to operate, I also investigated the role of landscape connectivity in explaining the distribution pattern of the species – either alone or used along with the spatial methods.

Together these four studies reveal the importance of many kinds of spatial effects controlling ecological processes, the effects of which are transferred to observed patterns. While relying on the theoretical background of classical spatial theories, my thesis aim at broadening the look on spatial effects by viewing landscapes as spatially structured and heterogeneous entities, and further complicating a scene by adding the hierarchical spatial effects. The important aim of the thesis is also to understand how the characteristics of the species themselves affect, in interaction with landscape structure, their responses. The thesis aims at demonstrating the independent and interrelated effects of the different spatial phenomena, and clarifies their relevance for species conservation.

3 MATERIALS AND METHODS

3.1 Boreal forest landscapes as a context

In the boreal forest zone, landscapes are characterized by a great share of forest cover. For example, in Finland 86 % of the land area is classified as forestry land (Anonymous 2010). However, within these seemingly forested landscapes many forest species suffer from the loss, degradation and fragmentation of habitats. The habitat loss is brought about by intensive modern forestry that has resulted in many kinds of changes in the structural features and compositions of forest stands. Forest stands that undergo typical forest management according to present-day forestry recommendations (including prescribed planting of seedlings, two to three high thinnings and clear-cutting after a growth period of ca. 80-100 years) (Anonymous 2007) are characterized by even-sized and -aged forest structure without natural vertical tree layering, lack of old deciduous trees and low amount of dead wood (Esseen et al. 1997, Östlund et al. 1997). At the larger scale, forest landscapes have been subject to changes in their composition and configuration; e.g. old-growth areas are increasingly separated by clear-cut areas and young stands (Kouki et al. 2001, Löfman and Kouki 2001).

According to latest IUCN evaluation (Rassi et al. 2010), a great share of threatened species in Finland (36 per cent) are dependent on forests. Forest species that are specialized in structural elements or habitats that are altered or decreased by forest management (e.g., coarse wood debris, productive forest biotopes, late successional stages) are the main sufferers. As the boreal forests are naturally very dynamic, many forest species have probably dispersal strategies well suited to locate available habitat and resources at the landscape level. In this kind of case, species respond mostly to the habitat loss caused by forest management. However, forest species with poor dispersal ability (for example, due to their dependence on static resources or stable habitats, see e.g., Hedin et al. 2008), or with otherwise restricted dispersal range (for example, due to social constraints), experience forested landscapes as dispersal-limited.

Then, in addition to habitat loss, species respond also to the effects of habitat fragmentation. Sensitivity to fragmentation intensifies the effects of habitat loss, and thereby represents one important species trait that inflates the vulnerability of a species to landscape level changes induced by commercial forestry (e.g., Penttilä et al. 2006).

3.2 Conservation area network analyses (I and II)

The network analyses conducted on reserve networks in three areas in Central Finland are based on digital datasets obtained from authorities, Metsähallitus and the forest company UPM Kymmene Oy. The reserve networks consisted of traditional conservation areas supplemented with woodland key habitats (WKHs). For traditional conservation areas, the data consisted of stand delineation with associated biotope information. For WKHs, the data provided information on the habitat delineation and the type(s) of the habitat. The three study areas comprised ca. 500 km^2 each, and they differed in terms of density of WKHs and coverage of conservation area network.

The network connectivity analyses were carried out using graph-theoretic connectivity approach. Graph-theory, or network theory, is a branch of science dedicated to the properties and functioning of networks (Gross and Yellen 2006). Graph-theory, with its supply of various algorithms, allows for analyses at the level of the whole network or at the level of individual patches. In landscape ecology, the network level analyses can be used to evaluate landscape connectivity (Urban and Keitt 2001, Fall et al. 2007, Urban et al. 2009). The analyses at the level of individual patches enable the connectivity contributions and functional roles of individual patches to be evaluated. The patch-level analyses may be based on network centrality measures (see e.g., Estrada & Bodin 2008) or so called node removal analysis. In the node removal analysis, each patch is systematically removed from the network and the reconstructed network is compared with the original network which included the patch (Keitt et al. 1997). The connectivity loss caused by the removal of a patch measures the contribution of the patch to the network connectivity. Patch-level connectivity evaluation with the mean of node removal analysis is often called the procedure of patch prioritization.

I varied a critical species attribute, i.e. dispersal ability, systematically from 200 m to 25 km. I constructed six habitat networks with a hierarchical formulation. Forests and peat lands were the habitat types highest in the hierarchy, and the other, more specialized habitat types were subsets of these more general networks. This resulted in habitat networks having a substantial variation in terms of patch number and network coverage. I compared the absolute and relative connectivity contributions from the WKHs by comparing the network comprised solely of traditional reserves to the reserve networks augmented with WKHs. Connectivity evaluations were based on Integral Index of Connectivity - measure, a binary index that allows patches to be linked also via indirect connections, i.e. paths (I). I also compared the connectivity evaluations of different graph theoretic connectivity measures, especially in terms of how they reacted when WKH patches were augmented to the reserve network (II). At the patch level, I assessed the connectivity contributions of individual WKH patches and their spatial potential to serve as a stepping stone patches in habitat networks (I). I evaluated also the consistency of different graph-theoretic connectivity measures to rate the connectivity values of habitat patches in three habitat networks varying in terms of patch density (II). As the connectivity value of a patch is combined effect of its area and position in a network, I evaluated the consistency separately for this composite value (i.e., determined by patch size and location) and for the value that is determined solely by patch location. I also investigated how different connectivity measures allocate relative weight between patch size and patch location on their connectivity ratings across the range of dispersal distances.

3.3 Distribution modeling of the Siberian Jay (III and IV)

The distribution modeling of the Siberian Jay was conducted in two areas (roughly 20 000 km^2 each), situated at the southern edge of the continuous range of the species. The modeling of the habitat association of the Siberian Jay (*Perisoreus infaustus*) was based on two different kinds of occurrence data. The one dataset consisted of sighting records (reported by ornithologists and public) from databases administered by local ornithological societies. The other dataset included occupancy information that was got from repeated feeding site surveys. The dataset from the feeding sites was burdened with imperfect detection so that any visit to an occupied site yielded detection with a probability of far less than one.

I used the multi-source Natural Forest Inventory (NFI9) data produced by The Finnish Forest Research Institute, National CORINE database of Finland and a raster-format digital elevation model. I calculated 14 territory level and 5 landscape level covariates (Fig. 1) as candidates to explain the species distribution pattern. I used an information theoretic framework to assess the support for candidate models formed as prescribed combinations of environmental covariates. I constructed two separate model sets; one included only territory level covariates (i.e., territory level model) and the other also landscape level covariates (i.e., multi-scale model). In the modeling, I accounted for the imperfect detection inherent in the occurrence dataset. As a preceding step of parametric modeling, I created additional pseudo-absence points to the improbable occurrence zone as predicted by the Environmental Niche Factor Analysis. This helped me to correct for the effect of survey effort that was biased towards high-quality sites, thus providing only limited range of variation for many environmental covariates.

For the best-supported models, I fitted Bayesian logistic regression and their correspondent spatially explicit Bayesian Gaussian Conditional autoregressive (CAR) models (III). In the Bayesian framework, I also accounted for the imperfect detection using a hierarchical model formulation. I evaluated the models based on their model fit, ability to remove residual autocorrelation and their discrimination capacity. For the best-performing model of Central Finland, I also fitted autologistic regression model to compare its parameter estimates and model performance with the ones produced by the non-spatial and CAR model (IV). Based on the predictions for the species occurrence probability, I also calculated a surface to represent landscape connectivity. This so-formed new variable was fed to modeling in order to investigate the role of landscape connectivity underlying the aggregation in species distribution (IV).

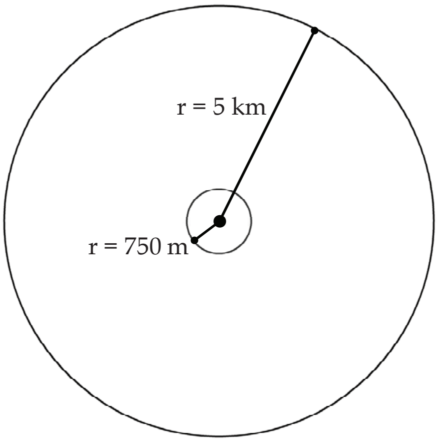


FIGURE 1 Territory and landscape level covariates were calculated with a circular moving window. A circle with a radius of 750 m encompasses area (1.8 km^2) of the order of territory size in Southern Finland, 1-5 km^2 (Lillandt 2000). A radius of 5 km corresponds to the average-range dispersal distances of the Siberian Jay (Gienapp & Merilä 2011).

4 RESULTS AND DISCUSSION

4.1 Connectivity contributions from the woodland key habitats (I)

My results show that the established reserve networks, with and without woodland key habitats (WKHs) set quite high demands on species dispersal ability in order to be functionally connected. Landscape connectivity seemed to be more informative measure for the amount of habitat actually available to organisms at the landscape level than total habitat area of a network, as the habitat area alone ignores the species specific differences in the access to available habitat. Species with poor dispersal ability were not able to exploit the habitat area provided by the WKHs to full degree. These species would require either larger set-asides or remarkably denser reserve network than exists presently. For the species with better dispersal ability, WKHs seemed to have potential to improve the network functionally by not only providing new protected habitat area but also by providing linkages between reserves.

For the rather rare habitat types (herb-rich forests and spruce mires), WKH-based conservation seemed to be an efficient way to improve network coverage and functional connectivity. These kinds of habitats are located also naturally highly scattered in a landscape, and a WKH-based, highly targeted conservation seems to be an efficient way to increase their representativeness in a reserve network.

Traditional reserves were better connected to other habitat patches than the WKH patches. These larger set-asides form the cornerstone of the reserve network, and form an essential support for the WKH-based conservation. WKH sites greatly benefit from having large reserve areas nearby. The results of the study show that also small-sized and scattered sites can be potentially valuable if the habitat distribution patterns and large-scale connectivity are acknowledged. Thus, spatial considerations do not exclusively support network aggregation and protection of large continuities, but may identify the combination of aggregation and scattering as a suitable conservation strategy.

The mere spatial properties of the WKHs are not a sufficient condition for their ecological value, but increased attention should be paid to maintaining their ecological quality in the future.

4.2 Insights into the evaluation of connectivity (II)

Although it is widely acknowledged that an increased connectivity is a desired property of habitat networks and individual patches, the definition and evaluation of connectivity is ambiguous. In this study, I compared the consistency and performance of various graph-theoretic connectivity measures in evaluating connectivity at the network and patch level. Graph-theoretic connectivity measures varied in terms of their conceptual underpinnings, and these differences were expectedly reflected to their ability to meaningfully measure network and patch level habitat connectivity.

Some measures expressed the addition of WKH patches to reserve network leading as a negative connectivity effect, especially for species with rather poor dispersal ability. All the measures that acknowledge the connectivity contributions between all the pairs of patches indicated a positive connectivity effect of adding the WKH patches. The measures also showed differences in terms of how the connectivity contributions from the WKH patches were rated as a function of dispersal ability. In contrast to the measures using a probabilistic connectivity model, measures with a binary connectivity model indicated an existence of a transition range for connectivity, i.e., a range where a small change in dispersal ability yields a huge increment in connectivity.

The patch prioritizations from the different connectivity measures were highly correlated, because the patch prioritizations were, to a large degree, driven by patch sizes. However, the prioritizations from the measures differed in terms of how much emphasis they put on patch size versus patch location across the range of dispersal distances. When the effect of patch size was controlled for, the measures value the locations of patches in a network context rather differently. Because in a sparser network the prioritizations were more determined by patch location than in a denser network, the inconsistency of patch prioritizations may be inflated in sparser networks.

My results underline the fact that the network connectivity analyses and patch prioritizations should be interpreted with caution. Different measures do not provide consistent solutions to the problems of connectivity management, but are operating within their conceptual restrictions. The performance of measures should be evaluated with a full range of empirical datasets so that their applicability would be better understood.

4.3 Habitat associations of the Siberian Jay in boreal forest landscapes (III)

While controlling for the imperfect detection and spatial autocorrelation, the Siberian Jay occurrence responded to the landscape context in both study areas. In the study area in central Finland, the Siberian Jay exhibited a stricter habitat association than in northern Karelia that is connected to a continuous distribution area of a northern Finland. In Central Finland, the occurrence of the Siberian Jay was especially enhanced by the spruce mires at the territory level and the forest cover at the landscape level. In Karelia, the occurrence was promoted, for example, by the mature spruce coverage at the territory level, and mires and low human impact at the landscape level. My results suggest that support the prevailing thinning recommendations may produce too open and uniform forest structure for the Siberian Jay.

The comparisons between the non-spatial logistic regression and spatial CAR models showed that the spatially explicit approach led to a better model fit and predictive capacity than did the non-spatial approach. Especially at the territory-level, the superiority of the spatially explicit modeling was striking. For the multi-scale models (i.e., models that incorporated both territory and landscape level covariates), the performance of the spatial models was not remarkably better compared with the non-spatial ones. This indicated that spatially structured error in the territory level models also accounted for the missing landscape level covariates. Based on the social system and restricted dispersal, it was to be expected that the species distribution was also aggregated due to intrinsic factors. This intrinsic aggregation was supported in the study area in Central Finland, where the spatially explicit approach further improved the model performance of the multi-scale models. It is also possible that the structure of landscape variables may partly and coincidentally account for the inherent structure of the response.

My results show that successful conservation for the species operates at the scales beyond the size of a single territory. For the long-term protection of the Siberian Jay, dense spruce forests and spruce mires should be available with suitable configuration at the landscape scale to support also the functioning of its spatially structured social system. This calls for aggregated conservation efforts in the landscape to enable natural rate of exchange of individuals between territories, and colonization of new empty territories near the source territories.

4.4 Alternative ways to account for space in distribution modeling (IV)

While the CAR model incorporates the spatial dependence in a response to a spatially structured error term, the autologistic regression puts the spatial dependency into a form of a new variable. My results showed that the parameter estimates from the CAR model and autologistic regression differed notably; the parameter estimates from the CAR model were constantly greater in magnitude than those from the autologistic model. Two parameters turned out to be insignificant when switched from the CAR model to the autologistic regression model. The parameter estimates from the CAR model were roughly of the same magnitude with the one produced by the nonspatial model, although the confidence intervals of the parameter estimates were wider for the CAR model than for the non-spatial model. The narrow confidence intervals of parameter estimates for the non-spatial model are indicative of the tendency of non-spatial models to overrate the certainty in parameter estimates, thus leading to too liberal test results.

However, when looked from the point of view of model outcome, modeling performance of the both spatial methods was superior to the one of the non-spatial model. Still, although their captured the same broad patterns for the species distribution, their predictions at the cell level showed inconsistencies. My results emphasize the fact that environment and space are strongly confounded, and modeling methods handle this interrelation between space and environment in different manners. Autologistic regression, that represents the spatial autocorrelation with a new variable, seemed to have greater tendency to allocate explanatory power to spatial effect (instead of environmental effect) than the CAR model, which in turn gives more weight to environmental factors. My results show that alternative modeling methods do not necessarily provide standard kind of control over spatial autocorrelation, but their characteristics can be also reflected to the inferences made about the association of species with their environments.

I investigated also the explanatory role of landscape connectivity in explaining species occurrence pattern. The addition of new variable representing landscape connectivity did improve the model performance compared to results of the non-spatial models, but the model performance lacked behind those of spatially explicit CAR models (with the second order neighborhood). The benefit of using connectivity information along with a CAR model was only marginal, which indicates that CAR model can successfully handle the spatial correlation rising from intrinsic aggregation. So, the CAR approach, incorporating both extrinsic and intrinsic factors behind species distributions, seemed to be efficient for predictive purposes. However, the enhanced model performance of the non-spatial model when supplemented with the landscape connectivity component provided a qualitative insight into a role of intrinsic autocorrelation shaping species distribution.

5 CONCLUSIONS AND IMPLICATIONS

Spatial structure in ecological patterns stems mostly from three reasons (see Wagner & Fortin 2005). First, species respond to environmental factors that exert control on their distribution and abundance, and those environmental factors themselves are spatially structured. The environment of a species is structured at many different scales, and with respect to many different factors at the same time. Second, species respond to their environments at their species-specific scales and in a hierarchical manner so that a complex interaction exists between the spatially structured environment and spatially structured species responses. And third, many ecological processes are spatial, controlled by distances between operating units, which leads to intrinsic autocorrelation in a response on top of the all extrinsically rising dependence. My results on the Siberian Jay manifest all these three mechanisms operating on the spatial patterning of its distribution.

Spatially structured ecological processes may bring about situations that species distributions and persistence cannot be understood and managed based on environmental factors or present-day species occurrences alone. For example, spatially structured processes may lead to species also occurring in suboptimal environmental conditions if supplied with the individuals from high-quality sources. Or alternatively, species may not occur in a high-quality site due to spatial constraints posed by a landscape. These both sides of a coin were demonstrated in the distribution pattern of the Siberian Jay in central Finland: occupied sites in the landscape strongly, and over the control of environmental factors, promoted the species occurrence in the neighboring sites. This effect was passed on to the species distribution pattern that expressed clear aggregation. Spatiality may break down the realization of the traditional niche concept, i.e. that species occur within the bounds of their environmental requirements in a n-dimensional environmental space, those bounds being also modified by biotic interactions (Hutchinson 1957). Thus, non-spatial snapshot views on distribution patterns may provide severely biased information on species niche and its relationship to environmental factors (see Pulliam 2000).

As indicated by my results on the connectivity of the reserve networks, the role of spatial planning plays a largest role for species that experience landscape as intermediately dispersal-limited. The greatest relative gainers of the WKH-based conservation (in terms of enhanced network connectivity) were the species with intermediate dispersal ability, for which the WKH sites efficiently linked otherwise unconnected protected areas. In cases where the dispersal is not a limiting factor, conservation can operate at the local scale without the need for rigorous spatial planning. On the other hand, species with highly restricted dispersal would mainly benefit from large-sized reserves, because for such species the landscape is too fragmented irrespective of additional connectivity provided by WKHs. Therefore, there exists a limited potential for spatial planning in fragmented landscapes for the poorly dispersing species. In a similar manner, also the landscape context affects the potential value of spatial conservation planning. There exists greatest potential for spatial planning in landscapes whose degree of fragmentation is intermediate.

The fact that the conservation value of a patch, along with other factors (such as its area, quality and species pool), is also determined by its spatial location in a habitat network, has important conservation implications. In practice, there might be situations where criteria for habitat quality, habitat area, recent species pool and spatial position of a proposed area in a network are in conflict. For example, when an isolated, high-quality proposed conservation area competes for reserve inclusion with the one of lower quality, well-connected area, there is a clear conflict. In order to weight these competing criteria in conservation decisions in an ecologically informed way, much knowledge on species-specific landscape responses is needed. As pointed out by my results on patch prioritizations derived from different connectivity measures, these measures do not provide consistent solutions to the weighting problem. Special emphasis should be given to the requirement that the modeled connectivity would really capture, or at least roughly approximate, the ecological process to be conserved.

The theoretical and empirical evidence clearly indicates the ecological benefits of aggregated conservation acts and reserve design. However, even for a single species, connectivity exists at many spatial scales. Connectivity one detects is always linked with the scale of the analyses. Conservation should safeguard landscape connectivity at all spatial scales. If conservation efforts are viewed and aggregated overly with a local emphasis, and the broader scale of connectivity is ignored, the processes operating over broad spatial domains may be compromised. Thus, even the best-planned, independently implemented conservation schemes with a rather local focus do not necessarily form a functional whole at the larger scales (like regional or national). Thus, the general recommendation of aggregating conservation effort in a landscape does not provide particularly unequivocal guideline for conservation, but should be accompanied by the explicit consideration of the spatial scales.

Conservation operates across the range of spatial scales – from global to very local ones. Connectivity conservation aims building interconnected systems of habitats. According to Bennet et al. (2006): "By increasing the flow of organisms or continuity of processes between parts of the landscape mosaics, there is potential to build habitat networks that integrate conservation efforts at multiple levels – including local, landscape, regional, continental and even global scales". Connectivity conservation may be thus viewed as a way to bridge these different spatial scales of conservations together, to form a functional whole in space and time. It is clear that the approaches and methods that give empirical support for the connectivity conservation also change with the scale of conservation (see Cabeza et al. 2010). When descended down from the broader spatial scales to local scales, general, multi-scale conservation problems are usually replaced by more specific ones. This is accompanied by the shift from rather implicit connectivity approaches to more explicit, complex and data-demanding, dynamic and species-centered ones.

Too often fragmentation is falsely equated with the indisputable changes in spatial patterning of landscapes brought about by habitat loss (increasing interpatch distances, smaller patch sizes etc.). Apart from the popular use of the concept of fragmentation, it is also the problem in scientific research that the two components, habitat loss and fragmentation, underlying the spatial patterning of habitat and affecting ecological processes remain coupled (Fahrig 2003). It would be important to make distinction between these two independent components when investigating species' landscape responses. This would make it possible to tell apart the real fragmentation sensitive species from the species that merely respond to the effects of habitat loss. This would bring along a more founded basis for spatial conservation planning.

In this thesis, I have worked upon a full range of spatial phenomena that underlie species distributions. My results have demonstrated that all these spatial effects are acting on species distribution patterns. Taking them into account in species conservation does not only safeguard the functioning of ecological processes, but at best brings about effectiveness into conservation. With careful spatial planning it is possible to yield a greater ecological effectiveness for a given investment to conservation. But to gain this effectiveness, the intricacies and complexities of spatial questions should be well understood both theoretically and empirically.

Acknowledgements

I have got through this process in the warm and expertised guidance of my two supervisors, Mikko and Janne. Mikko, I could not imagine a more suitable supervisor for me than you have been. You have given me enough space to breathe and express myself, but still I have felt your caring guidance all the time. You have always confronted me in a very warm and human way, and I have always felt that you see me as a whole person. Janne, I really see you as a great advocate of scientific values and thinking. I admire your openness to new ideas, and I hope that I could possess even a small piece of your scientific attitude. All the time, I have experienced you as a safe cornerstone, a person that I can really account on. You are like two little boys, getting constantly excited about the scientific questions and complexities of life.

Merja and Jonna, you have been like my scientific sisters during these years – thank you for your warm support and memorable moments together. Mikael, you have a special ability to cheer me up and laugh at silly and simple things with me- I appreciate you as a friend and a great scientist. For everyday laughs and companion, I would especially like to thank, all equally, Aapo, Atte, Elisa, Inka, Katja and Panu – life would be so much duller without you, dears. Nina, thank you for your companionship during our shared APO-year and for the nice time I have spent with you. Life wouldn't be the same without the MCC – thank you all, former and present club members, for your unfiltered thoughts, grass-root level empathy and nice moments together. I have been lucky to share my working room with many great persons; Saija, Sanna, Jenni, Jouni, Tuomas, Janne, Santtu, Adriano, and now finally, Hannele and Anssi. These roomacquaintances have led to me, for example, to the world of whiskies, moose hunting, toilet renovations and overwhelming influences of female hormones, as well as visits to the newest night bars in the city. Pasi, I have experienced you as my closest colleague sharing my interest in all kind of spatial issues– thank you for your inspirational effect on me and your fresh and broad-minded thoughts about the science and life in general. Kristiina, thank you for sharing your interest on the Siberian Jays with me. I have learned much from you, even though I should be the one that is supervising. Pete A., you have had a long-lasting influence on me – you believed in my capacities even at the early stage of my basic studies and gave opportunities to familiarize myself with many kinds of scientific issues. This all has provided me a mental resource that has turned out to be important in my life and career ever since. Markku, you have provided me many valuable opportunities along my career – for me, you represent the human side in the scientific world. I would like to thank all my colleagues at the department – you all are collectively behind the nice atmosphere there!

My life is blessed with many dear friends and relatives, who have also played a great role in this thesis. Families Laita (in Espoo), Myllys, Penttinen and Selonen – warm thanks for the long-term participation in my life. Annukka especially, thank you for your heart-warming support flavored with a full range of feelings. My friends Irma and Katja – you always remind me of our shared life as students, and luckily our friendship has also tracked the passage of time. I have also many important friends that I have met via kids, most of all I would like to thank families Mikkonen and Leino. Dear wall-neighbors, family Tonteri, thank you for your valuable help and warm support in those times when we have run out of capabilities or time. Tarja, my mother-in-law, you have offered invaluable practical help during our occupied years. It has been luxury to enter to the cleaned house after your visits. My sister Tuija and your whole family, calling to you has always produced me a feeling that our life is not that chaotic after all – you always have wilder moments going on there. Äiti, even though I am afraid that you are not that interested in my scientific issues, you have participated in my life valuably in all other respects. You have not once pushed me according to your own interests – I really appreciate I have allowed to find an uninfluenced path of my own.

Lluís and Miska, great thanks for reviewing and commenting on my thesis. Jari, I am grateful for your sharp-eyed editorial work.

For funding of my work, I would like to thank Kone Foundation, the Rector of University of Jyväskylä and the cluster programme "Talousmetsät, pienialaiset suojelukohteet ja suojelualueet monimuotoisuuden suojelussa: Integroitu ekologinen vaikuttavuus ja kustannustehokkuus".

My furry friend Macho - you have been my mental healer and more important for me than anyone could imagine. Pihla and Elsi, with you in my life I always stay touched to real life and my ultimate values. I think that your imaginations have nurtured also my scientific work. Mika, these years have been quite heavy, but still our love has carried us through the difficulties. Thank you for your unshakeable support and faith in me. My dear family, with your loving support (and in spite of you all), I am now about to bring this project to a closure. And a closure is always a launch for a something new.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Lajien suojelu tilassa

Lajien levinneisyys maisematasolla on useiden tilajärjestykseen nivoutuvien ilmiöiden, prosessien ja niiden välisten vuorovaikutusten monimutkainen lopputulos. Lajeja säätelevät elinympäristötekijät ovat tilallisesti rakentuneita, jolloin niiden tilarakenne välittyy myös lajien esiintymiseen. Lisäksi lajit vastaavat elinympäristöönsä yhtäaikaisesti usealla eri tilallisella mittakaavalla. Lajien ekologiset ominaisuudet vaikuttavat näihin hyvin lajikohtaisten tilallisten mittakaavojen määräytymiseen. Tämän kaiken lisäksi ekologiset prosessit itsessään ovat tilajärjestykseen sitoutuneita, mikä aiheuttaa lajien esiintymiseen yleensä vielä elinympäristön tilajärjestystä voimakkaamman tilallisen rakenteen.

Tilaan liittyvät ekologiset ilmiöt korostuvat sopivan elinympäristön vähentyessä maisematasolla. Tämän vuoksi tilaan liittyvien suojelukysymysten merkitys kasvaa erityisesti uhanalaisilla lajeilla, joiden elinympäristöjä ihmistoiminta on voimakkaasti vähentänyt ja heikentänyt. Tilassa tapahtuvien ekologisten prosessien ymmärtäminen ja huomioiminen on tärkeä osa lajistonsuojelua, jonka tavoitteena on lajien pitkäaikainen säilyttäminen muuttuvassa maailmassa. Lisäksi tilallisten tekijöiden huomioimisella on tärkeä osa ekologisen tiedon käsittelyssä, jotta lajien esiintymisen ja runsauden suhteesta niitä sääteleviin elinympäristötekijöihin voidaan tehdä oikeita johtopäätöksiä.

Väitöskirjani yhtenä boreaalisten metsien suojelukysymyksenä tutkin metsälakikohteiden merkitystä ja roolia osana perinteisten suojelualueiden muodostamaa maisematason suojeluverkostoa. Käytin arvioinnissa verkkoteoriaa, joka tarjosi mahdollisuuksia tarkastella metsälakikohteiden merkitystä niin koko verkoston kuin yksittäisten suojelukohteiden tasolla. Tulokset osoittivat, että lajien liikkumiskyvyllä voi olla merkittävä vaikutus metsälakikohteiden tarjoamaan suojeluhyötyyn. Kytkeytyneisyysarviot voivat antaa realistisemman kuvan suojeltavan pinta-alan ekologisesta vaikuttavuudesta kuin itse pinta-ala. Vaikka suuret suojelualueet ovat luonnonsuojelun tukiranka, maisematason kytkeytyneisyyttä voidaan suhteellisen hyvin liikkuvien lajien osalta parantaa myös pienialaisilla, talousmetsien suojelukohteilla. Erityisen tehokkaasti metsälakikohteet lisäsivät verkostossa harvinaisten ja rehevien biotooppien (esimerkkeinä lehdot ja rehevät korvet) edustusta, koska näiden biotooppien luontaisesti laikuttainen esiintymiskuva on vaikea saada turvattua suurten, keskittyneiden suojelualueiden keinoin.

Toisena väitöskirjan suojelukysymyksenä tutkin lajistonsuojelua ja siihen liittyviä tilakysymyksiä kuukkelin (*Perisoreus infaustus*) näkökulmasta. Kuukkeli on eteläsuomalaisessa metsämaisemassa viime vuosikymmeninä voimakkaasti taantunut ja nykyään hyvin laikuttaisesti esiintyvä pohjoisten taigametsien lintulaji. Tutkin kuukkelin elinympäristövaatimuksia eteläisen Suomen metsämaisemassa levinneisyysmallinnuksen keinoin. Mallinnuksessa käsittelin kuukkelin elinympäristösuhdetta kahdella tilallisella mittakaavalla, reviiritasolla ja maisematasolla. Käsittelin mallinnuksessa niin ekologisista prosesseista kuin elinympäristön rakenteellisuudesta välittyvää vaikutusta kuukkelin esiintymisen jakautumiseen maisematasolla. Tilallisten tekijöiden eksplisiittinen huomioiminen mallinnuksessa tarjosi yhä paremman ja luotettavamman mahdollisuuden ymmärtää lajin elinympäristösuhdetta metsämaisematasolla. Tulokset antoivat viitettä siitä, että lajin esiintymistä metsämaisemassa yhtenäisen levinneisyysalueen ulkopuolella on vaikea ymmärtää puhtaasti elinympäristömuuttujien perusteella, vaan lajin ekologiset ominaisuudet ja käyttäytymispiirteet vaikuttavat sen esiintymiskuvaan. Lisäksi myös maisematason tekijät selittivät tuloksissani kuukkelin esiintymistä. Tulokseni antavat viitettä, että lajin suojelua täytyy tarkastella tarpeeksi suurella mittakaavalla, ja suojeluponnistuksia pitäisi keskittää maisematasolla.

Yksi keskeinen osa väitöskirjaani oli myös erilaisten tilaan liittyvien menetelmien (kytkeytyneisyysmittareiden ja tilallisten mallinnusmenetelmien) vertailu. Tulokset osoittivat, että tilalliset menetelmät itsessään voivat vaikuttaa tuloksiin. Tämän vuoksi tilallisten analyysien tulkinta ja soveltaminen vaatii laajaa ymmärtämistä niiden takana olevista oletuksista ja tilaan liittyvien kysymysten monimuotoisuudesta.

Väitöskirjani tuo yhteen monia tilallisia ilmiöitä ja niiden käsittelyyn tarkoitettuja menetelmiä. Metsämaisemataso on monien ekologisten prosessien näyttämö, jossa tapahtuvat sekä elinympäristön ja eliöiden että eliöiden väliset vuorovaikutukset. Monilla näillä prosesseilla on tilallinen ulottuvuus. Tämän ulottuvuuden yhä monipuolisempi huomioiminen tarjoaa mahdollisuuksia ymmärtää paremmin lajien esiintymistä sääteleviä tekijöitä ja prosesseja pirstoutuneissa metsämaisemissa. Tätä ymmärrystä voidaan käyttää hyvin monella tavalla hyödyksi käytännön lajistonsuojelutyössä.

REFERENCES

Anonymous 2007. *Hyvän metsänhoidon suositukset.* Metsäkustannus, Helsinki.

Anonymous 2010. *Finnish statistical yearbook of forestry.* Metsäntutkimuslaitos, Helsinki.

Andrén H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: A review. *Oikos* 71: 355-366.

Angelstam, P., Boutin, S., Schmiegelow, F., Villard, M.-A., Drapeau, P., Host, G., Innes,J., Isachenko, G., Kuuluvainen, T., Mönkkönen, M., Niemelä, J., Niemi, G., Roberge,J.-M., Spence, J. and Stone, D. 2004. Targets for boreal forest biodiversity conservation– a rationale for macroecological research and adaptive management. – *Ecol Bull* 51:487–509.

Araújo M.B. & Williams P.H. 2000. Selecting areas for species persistence using occurrence data. *Biol Conserv* 96: 331-345.

Augustin N.H., Mugglestone M.A. & Buckland S.T. 1996. An Autologistic Model for the Spatial Distribution of Wildlife. *J Appl Ecol* 33: 339-347.

Austin M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157: 101-118.

Bahn V., Krohn W.B. & O'Connor R.J. 2008. Dispersal leads to spatial autocorrelation in species distributions: A simulation model. *Ecol Model* 213: 285-292.

Bakker K.K., Naugle D.E. & Higgins K.F. 2002. Incorporating landscape attributes into models for migratory grassland bird conservation. *Conserv Biol* 16: 1638-1646.

Bender D.J. & Fahrig L. 2005. Matrix structure obscures the relationship between interpatch movement and patch size and isolation. *Ecology* 86: 1023-1033.

Bennet, A.E., Crooks, K.R. & Sanjayan, A. 2006. The future of connectivity conservation. In: Crooks K.R. & Sanjayan M. (Eds.) 2006. *Connectivity conservation.* Cambridge University Press, Cambridge., pp. 676-694.

Bini L.M., Diniz-Filho J.A.F., Rangel T.F.L.V.B., Akre T.S.B., Albaladejo R.G., Albuquerque F.S., Aparicio A., Araújo M.B., Baselga A., Beck J., Bellocq M.I., Böhning-Gaese K., Borges P.A.V., Castro-Parga I., Chey V.K., Chown S.L., Marco P.d., Jr, Dobkin D.S., Ferrer-Castán D., Field R., Filloy J., Fleishman E., Gómez J.F., Hortal J., Iverson J.B., Kerr J.T., Kissling W.D., Kitching I.J., León-Cortés J.L., Lobo J.M., Montoya D., Morales-Castilla I., Moreno J.C., Oberdorff T., Olalla-Tárraga M.Á., Pausas J.G., Qian H., Rahbek C., Rodríguez M.Á., Rueda M., Ruggiero A., Sackmann P., Sanders N.J., Terribile L.C., Vetaas O.R. & Hawkins B.A. 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography* 32: 193-204.

Brouwers N.C. & Newton A.C. 2009a. Movement rates of woodland invertebrates: a systematic review of empirical evidence. *Insect Conservation and Diversity* 2: 10-22.

Brouwers N. & Newton A. 2009b. The influence of habitat availability and landscape structure on the distribution of wood cricket (Nemobius sylvestris) on the Isle of Wight, UK. *Landscape Ecol* 24 . 199-212.

Brückmann S.V., Krauss J. & Steffan-Dewenter I. 2010. Butterfly and plant specialists suffer from reduced connectivity in fragmented landscapes. *J Appl Ecol* 47: 799-809.

Buler J.J., Moore F.R. & Woltmann S. 2007. A multi-scale examination of stopover habitat use by birds. *Ecology* 88: 1789-1802.

Cabeza M. 2003. Habitat loss and connectivity of reserve networks in probability approaches to reserve design. *Ecol Lett* 6: 665-672.

Cabeza M. & Moilanen A. 2003. Site-Selection Algorithms and Habitat Loss. *Conserv Biol* 17: 1402-1413.

Cabeza M., Arponen A., Jäättelä L., Kujala H., Van Teeffelen A. & Hanski I. 2010. Conservation planning with insects at three different spatial scales. *Ecography* 33: 54-63.

Chardon J., Adriaensen F. & Matthysen E. 2003. Incorporating landscape elements into a connectivity measure: a case study for the Speckled wood butterfly (Pararge aegeria L.). *Landscape Ecol* 18: 561-573.

Cronin J.T. 2007. From population sources to sieves: the matrix alters host-parasitoid source-sink structure. *Ecology* 88: 2966-2976.

Crooks K.R. & Sanjayan M. (Eds.) 2006. *Connectivity conservation.* Cambridge University Press, Cambridge.

Cushman S.A. & McGarigal K. 2004. Hierarchical analysis of forest bird species-Environment relationships in the Oregon Coast Range. *Ecol Appl* 14: 1090-1105.

Dale M.R.T. & Fortin M.J. 2002. Spatial autocorrelation and statistical tests in ecology. *Ecoscience* 9: 162-167.

Diamond J.M. 1975. The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biol Conserv* 7: 129-146.

Diniz-Filho J.A.F., Bini L.M. & Hawkins B.A. 2003. Spatial autocorrelation and red herrings in geographical ecology. *Global Ecol Biogeogr* 12: 53-64.

Diniz-Filho J.A.F., Rangel,Thiago Fernando L. V. B. & Bini L.M. 2008. Model selection and information theory in geographical ecology. *Global Ecol Biogeogr* 17: 479-488.

Dormann C.F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecol Biogeogr* 16: 129-138.

Dormann C.F., McPherson J.M., Araujo M.B., Bivand R., Bolliger J., Carl G., Davies R.G., Hirzel A., Jetz W., Kissling W.D., Kuhn I., Ohlemuller R., Peres-Neto P.R., Reineking B., Schroder B., Schurr F.M. & Wilson R. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30: 609-628.

Drapeau P., Leduc A., Giroux J., Savard J.L., Bergeron Y. & Vickery W.L. 2000. Landscape-scale disturbances and changes in bird communities of boreal mixed-wood forests. *Ecol Monogr* 70: 423-444.

Dungan J.L., Perry J.N., Dale M.R.T., Legendre P., Citron-Pousty S., Fortin M., Jakomulska A., Miriti M. & Rosenberg M.S. 2002. A balanced view of scale in spatial statistical analysis. *Ecography* 25: 626-640.

Edman M., Kruys N. & Jonsson B.G. 2004a. Local dispersal sources strongly affect colonization patterns of wood-decaying fungi on spruce logs. *Ecol Appl* 14: 893-901.

Edman M., Gustafsson M., Stenlid J., Jonsson B.G. & Ericson L. 2004b. Spore deposition of wood-decaying fungi: importance of landscape composition. *Ecography* 27: 103-111.

Elith J. & Leathwick J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40: 677-697.

Estrada E. & Bodin O. 2008. Using network centrality measures to manage landscape connectivity. *Ecol Appl* 18:1810-1825.

Esseen P., Ehnström B., Ericsson L. & Sjöberg K. 1997. Boreal forests. *Ecol Bull* : 16-47.

Fahrig L. 1997. Relative effects of habitat loss and fragmentation on population extinction. *J Wildl Manage* 61: 603-610.

Fahrig L. 2001. How much habitat is enough? *Biol Conserv* 100: 65-74.

Fahrig L. 2003. Effects of habitat fragmentation on biodiversity. *Annu Rev Ecol Evol Syst* 34: 487-515.

Fall A., Fortin M., Manseau M. & O'Brien D. 2007. Spatial graphs: principles and applications for habitat donnectivity. *Ecosystems* 10: 448-461.

Fischer J. & Lindenmayer D.B. 2007. Landscape modification and habitat fragmentation: a synthesis. *Global Ecol Biogeogr* 16: 265-280.

Fortin M. & Dale M. 2005. *Spatial analysis. A guide for ecologists.* Cambridge University Press.

Franken R.J. & Hik D.S. 2004. Influence of habitat quality, patch size and connectivity on colonization and extinction dynamics of collared pikas Ochotona collaris. *J Anim Ecol* 73: 889-896.

Gienapp P. & Merilä J. 2011. Sex-specific fitness consequences of dispersal in Siberian jays. *Behav Ecol Sociobiol* 65: 131-140.

Gross J.L. & Yellen J. 2006. *Graph theory and its applications.* Chapman & Hall, CRC, Boca Raton.

Guisan A. & Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8: 993-1009.

Guisan A. & Zimmermann N.E. 2000. Predictive habitat distribution models in ecology. *Ecol Model* 135: 147-186.

Gustafson E.J. & Gardner R.H. 1996. The effect of landscape heterogeneity on the probability of patch colonization. *Ecology* 77: 94-107.

Haila Y. 2002. A conceptual genealogy of fragmentation research: From island biogeography to landscape ecology. *Ecol Appl* 12: 321-334.

Hanski I. 1994. A Practical Model of Metapopulation Dynamics. *J Anim Ecol* 63: 151-162.

Hanski I. 1998. Metapopulation dynamics. *Nature* 396: 41-49.

Hanski I. 1999. *Metapopulation Ecology.* Oxford University Press, Oxford, UK.

Hanski I., Kuussaari M. & Nieminen M. 1994. Metapopulation structure and migration in the butterfly Melitaea Cinxia. *Ecology* 75: 747-762.

Hansson L. 2001. Key habitats in Swedish managed forests. S*cand J Forest Res 3* (Suppl.): 52-61.

Hedin J., Ranius T., Nilsson S. & Smith H. 2008. Restricted dispersal in a flying beetle assessed by telemetry. *Biodivers Conserv* 17: 675-684.

Henle K., Davies K.F., Kleyer M., Margules C. & Settele J. 2004. Predictors of species sensitivity to fragmentation. *Biodivers Conserv* 13: 207-251.

Hill J.K., Thomas C.D. & Huntley B. 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. *Proc R Soc Lond [Biol]* 266: 1197-1206.

Hill J.K., Collingham Y.C., Thomas C.D., Blakeley D.S., Fox R., Moss D. & Huntley B. 2001. Impacts of landscape structure on butterfly range expansion. *Ecol Lett* 4: 313-321.

Hoeting J.A. 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecol Appl* 19: 574-577.

Hoeting J.A., Davis R.A., Merton A.A. & Thompson S.E. 2006. Model selection for geostatistical models. *Ecol Appl* 16: 87-98.

Hutchinson G. 1957. Population studies - Animal ecology and memography - Concluding remarks. *Cold Spring Harb Symp Quant Biol* 22: 415-427.

Ishihama F., Takeda T., Oguma H. & Takenaka A. 2010. Comparison of effects of spatial autocorrelation on distribution predictions of four rare plant species in the Watarase wetland. *Ecol Res* 25: 1057-1069.

Keitt T., Urban D., & Milne B. 1997. Detecting critical scales in fragmented landscapes. *Conservation Ecology* 1:4

Keitt T.H., Bjornstad O.N., Dixon P.M. & Citron-Pousty S. 2002. Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25: 616-625.

Kindlmann P. & Burel F. 2008. Connectivity measures: a review. *Landscape Ecol* 23: 879-890.

Knapp R.A., Matthews K.R., Preisler H.K. & Jellison R. 2003. Developing probabilistic models to predict amphibian site occupancy in a patchy landscape. *Ecol Appl* 13: 1069-1082.

Kouki J., Löfman S., Martikainen P., Rouvinen S. & Uotila A. 2001. Forest fragmentation in Fennoscandia:linking habitat requirements of wood-associated threatened species to landscape and habitat changes. *Scand J Forest Res* 16: 27.

Kühn I. 2007. Incorporating spatial autocorrelation may invert observed patterns. *Divers Distrib* 13: 66-69.

Legendre P. 1993. Spatial Autocorrelation - Trouble Or New Paradigm. *Ecology* 74: 1659-1673.

Levin S.A. 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. *Ecology* 73: 1943-1967.

Lichstein J.W., Simons T.R., Shriner S.A. & Franzreb K.E. 2002b. Spatial autocorrelation and autoregressive models in ecology. *Ecol Monogr* 72: 445-463.

Lillandt B. G. 2000. Suupohjan kuukkelitutkimus 27 vuotta 1974-2000. *Hippiäinen* 30: 13-26.

Lindenmayer D. & Franklin J.F. 2002. *Conserving forest biodiversity : a comprehensive multiscaled approach.* Island Press, Washington.

Löfman S. & Kouki J. 2001. Fifty years of landscape transformation in managed forests of southern Finland. *Scand J For Res* 16: 44.

Manning A.D., Lindenmayer D.B. & Nix H.A. 2004. Continua and Umwelt: novel perspectives on viewing landscapes. *Oikos* 104: 621-628.

MacArthur, R. H. and Wilson, E. O. 1967. *The Theory of Island Biogeography*. Princeton, Princeton University Press.

Margules C.R. & Pressey R.L. 2000. Systematic conservation planning. *Nature* 405: 243-253.

Matisziw T. & Murray A. 2009. Connectivity change in habitat networks. *Landscape Ecol* 24 . 89-100.

Melles S.J., Fortin M., Lindsay K. & Badzinski D. 2011. Expanding northward: influence of climate change, forest connectivity, and population processes on a threatened species' range shift. *Global Change Biol* 17: 17-31.

Merriam, G. 1984. Connectivity: a fundamental ecological characteristic of landscape patterns. *Proceedings of the International Association for Landscape Ecology* 1: 5-15.

Miller J., Franklin J. & Aspinall R. 2007. Incorporating spatial dependence in predictive vegetation models. *Ecol Model* 202: 225-242.

Moilanen A. & Hanski I. 2001. On the use of connectivity measures in spatial ecology. *Oikos* 95: 147-151.

Mönkkönen M. & Reunanen P. 1999. On critical thresholds in landscape connectivity: a management perspective. *Oikos* 84: 302-305.

Nilsson, S.G. 2009. Selecting biodiversity indicators to set conservation targets: species, structures or processess. In: Villard M. & Jonsson B.G. (eds.) 2009. *Setting conservation targets for managed forest landscapes.* Cambridge University Press, Cambridge, UK., pp. 79-108.

Öckenger E., Niklasson M. & Nilsson S.G. 2005. Is local distribution of the epiphytic lichen Lobaria pulmonaria limited by dispersal capacity or habitat quality? *Biodivers Conserv* 14: 759-773.

Önal H. & Briers R. 2005. Designing a conservation reserve network with minimal fragmentation: A linear integer programming approach. *Environ Model Assess* 10: 193-202.

Opdam P. & Wascher D. 2004. Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. *Biol Conserv* 117: 285-297.

Östlund L., Zackrisson O. & Axelsson A.L. 1997. The history and transformation of a Scandinavian boreal forest landscape since the 19th century. *Can J Forest Res* 27: 1198-1206.

Pearson R.G. & Dawson T.P. 2005. Long-distance plant dispersal and habitat fragmentation: identifying conservation targets for spatial landscape planning under climate change. *Biol Conserv* 123: 389-401.

Penttilä R., Lindgren M., Miettinen O., Rita H. & Hanski I. 2006. Consequences of forest fragmentation for polyporous fungi at two spatial scales. *Oikos* 114: 225-240.

Prevedello J. & Vieira M. 2010. Does the type of matrix matter? A quantitative review of the evidence. *Biodivers Conserv* 19: 1205-1223.

Prugh L.R., Hodges K.E., Sinclair A.R.E. & Brashares J.S. 2008. Effect of habitat area and isolation on fragmented animal populations. *Proc Natl Acad Sci USA* 105: 20770-20775.

Pulliam H.R. 1988. Sources, sinks, and population regulation. *Am Nat* 132: 652-661.

Pulliam H.R. 2000. On the relationship between niche and distribution. *Ecol Lett* 3: 349-361.

Ranius T. 2006. Measuring the dispersal of saproxylic insects: a key characteristic for their conservation. *Popul Ecol* 48: 177-188.

Ranius T., Johansson V. & Fahrig L. 2010. A comparison of patch connectivity measures using data on invertebrates in hollow oaks. *Ecography* 33: 971-978.

Rassi, P., Hyvärinen, E., Juslén, A. and Mannerkoksi, I. (Eds.) 2010. The 2000 Red List of Finnish Species. Ympäristöministeriö, Suomen Ympäristökeskus, Helsinki (in Finnish with English summary)

Rayfield B., Moilanen A. & Fortin M. 2009. Incorporating consumer–resource spatial interactions in reserve design. *Ecol Model* 220: 725-733.

Ricketts T.H. 2001. The matrix matters: effective isolation in fragmented landscapes. *Am Nat* 158: 87-99.

Santika T. & Hutchinson M.F. 2009. The effect of species response form on species distribution model prediction and inference. *Ecol Model* 220: 2365-2379.

Schleicher A., Biedermann R. & Kleyer M. 2011. Dispersal traits determine plant response to habitat connectivity in an urban landscape. *Landscape Ecol* 26: 529-540.

Schneider D.C. 2001. The rise of the concept of scale in ecology. *Bioscience* 51: 545-553.

Segelbacher G., Cushman S., Epperson B., Fortin M., Francois O., Hardy O., Holderegger R., Taberlet P., Waits L. & Manel S. 2010. Applications of landscape genetics in conservation biology: concepts and challenges. *Conserv Genet* 11: 375-385.

Segurado P., Araújo M.B. & Kunin W.E. 2006. Consequences of spatial autocorrelation for niche-based models. *J Appl Ecol* 43: 433-444.

Sirami C., Brotons L. & Martin J. 2008. Spatial extent of bird species response to landscape changes: colonisation/extinction dynamics at the community-level in two contrasting habitats. *Ecography* 31: 509-518.

Spring D., Baum J., Nally R.M., Mackenzie M., Sanchez-Azofeifa A. & Thomson J.R. 2010. Building a regionally connected reserve network in a changing and uncertain world. *Conserv Biol* 24: 691-700.

Taylor P.D., Fahrig L., Henein K. & Merriam G. 1993. Connectivity is a vital element of landscape structure. *Oikos* 68: 571-573.

Taylor, P.D:, Fahrig, L., With, K.A. 2006. Landscape connectivity: a return to the basics. In: Crooks K.R. & Sanjayan M. (Eds.) *Connectivity conservation.* Cambridge University Press, Cambridge, pp. 29-43.

Teeffelen A.J. & Ovaskainen O. 2007. Can the cause of aggregation be inferred from species distributions? *Oikos* 116: 4-16.

Tischendorf L. & Fahrig L. 2000. On the usage and measurement of landscape connectivity. *Oikos* 90: 7-19.

Tischendorf L. & Fahrig L. 2001. On the use of connectivity measures in spatial ecology. A reply. *Oikos* 95: 152-155.

Timonen J., Siitonen J., Gustafsson L., Kotiaho J.S., Stokland J.N., Sverdrup-Thygeson A. & Mönkkönen M. 2010. Woodland key habitats in northern Europe: concepts, inventory and protection. *Scand J For Res* 25: 309-324.

Tognelli M.F. & Kelt D.A. 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography* 27: 427-436.

Trzcinski M.K., Fahrig L. & Merriam G. 1999. Independent Effects of Forest Cover and Fragmentation on the Distribution of Forest Breeding Birds. *Ecol Appl* 9: 586-593.

Turner M.G., Gardner R.H. & O'Neill R.V. 2001. *Landscape ecology in theory and practice : pattern and process.* Springer, New York.

Urban D. & Keitt T. 2001. Landscape Connectivity: A Graph-Theoretic Perspective. *Ecology* 82: 1205-1218.

Urban D.L., Minor E.S., Treml E.A. & Schick R.S. 2009. Graph models of habitat mosaics. *Ecol Lett* 12: 260-273.

Väisänen R.A., Lammi E., Koskimies P. & Kostet J. 1998. *Muuttuva pesimälinnusto.* Otava, Helsinki.

van Langevelde F. 2000. Scale of habitat connectivity and colonization in fragmented nuthatch populations. *Ecography* 23: 614-622.

van Teeffelen A.J.A, Cabeza M. & Moilanen A. 2006. Connectivity, probabilities and persistence: comparing reserve selection strategies. *Biodivers Conserv* 15: 899-919.

Villard M. & Jonsson B.G. (Eds.) 2009. *Setting conservation targets for managed forest landscapes.* Cambridge University Press, Cambridge, UK.

Villard M., Trzcinski M.K. & Merriam G. 1999. Fragmentation effects on forest birds: relative influence of woodland cover and configuration on landscape occupancy. *Conserv Biol* 13: 774-783.

Wagner H.H. & Fortin M. 2005. Spatial analysis of landscapes: concepts and statistics. *Ecology* 86: 1975-1987.

Watling J.I., Nowakowski A.J., Donnelly M.A. & Orrock J.L. 2011. Meta-analysis reveals the importance of matrix composition for animals in fragmented habitat. *Global Ecol Biogeogr* 20: 209-217.

Wiegand T., Revilla E. & Moloney K.A. 2005. Effects of habitat loss and fragmentation on population dynamics. *Conserv Biol* 19: 108-121.

Wiens J.A. 1989. Spatial scaling in ecology. *Funct Ecol* 3: 385-397.

Williams J., ReVelle C. & Levin S. 2005. Spatial attributes and reserve design models: A review. *Environ Model Assess* 10: 163-181.

Williams P., Hannah L., Andelman S., Midgley G., Araújo M., Hughes G., Manne L., Martinez-Meyer E. & Pearson R. 2005. Planning for climate change: identifying minimum-dispersal corridors for the Cape Proteaceae. *Conserv Biol* 19: 1063-1074.

Wintle B.A. & Bardos D.C. 2006. Modeling species-habitat relationshipts with spatially autocorrelated observation data. *Ecol Appl* 16: 1945-1958.

With K.A. & King A.W. 1999. Extinction thresholds for species in fractal landscapes. *Conserv Biol* 13: 314-326.

ORIGINAL PAPERS

I

WOODLAND KEY HABITATS EVALUATED AS PART OF A FUNCTIONAL RESERVE NETWORK

by

Anne Laita, Mikko Mönkkönen & Janne S. Kotiaho 2010

Biological Conservation 143: 1212-1227.

Biological Conservation 143 (2010) 1212–1227

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Woodland key habitats evaluated as part of a functional reserve network

Anne Laita *, Mikko Mönkkönen, Janne S. Kotiaho

Department of Biological and Environmental Science, University of Jyväskylä, PO Box 35, Survontie 9, FIN-40014, Finland

ARTICLE INFO

Article history:
Received 25 June 2009
Received in revised form 21 January 2010
Accepted 21 February 2010
Available online 16 March 2010

Keywords:
Woodland key habitat
Graph theory
Reserve network
Connectivity
Centrality measure
Integral Index of Connectivity

ABSTRACT

Woodland key habitats (WKHs) represent a potentially cost-efficient means to protect biodiversity in managed forests. The Forest Act of Finland defines 13 habitat types of WKHs, which enjoy legal protection. It has been argued that WKHs are too small-sized and scattered in occurrence to be actually important in the maintenance of forest biodiversity. However, from the species' perspective, WKHs form a network together with nature reserves. We evaluated the value and role of WKHs as a part of the whole reserve network using a graph-theoretical connectivity approach in three areas (ca. 500 km^2 each) located in Central Finland. The networks were formed separately for different habitat types and dispersal distances (ranging from 200 m to 25 km). We compared networks with and without WKHs, and thereby quantified the contribution of WKHs to overall network connectivity. We also examined the role of WKHs in the networks based on patch importance and network centrality measures. The results showed that the connectivity contributions of WKHs are tightly linked with the dispersal abilities of threatened species: WKHs enhance habitat connectivity, especially for species with an intermediate dispersal ability. For species with a poor dispersal ability, the protection of large set-asides would be a more efficient way to increase habitat connectivity than WKHs. WKH-based conservation seems to improve the connectivity of naturally rare and scattered habitat types relatively more than common habitat types, but in sparse networks a greater dispersal ability is required to gain enhanced connectivity than in dense ones. The connectivity value of WKHs can be understood as an emergent and scale-dependent property, appearing at the level of the entire functional network. Provided that the site characteristics of WKHs can be safeguarded, they can be a valuable and efficient addition to the reserve network.

1. Introduction

Habitat loss and fragmentation are major global threats to biodiversity, and efficient tools are needed to combat their negative effects. Supplementing existing protected-area networks is one of these tools. The history of protected-area establishment in many global regions has produced a network that is biased toward infertile landscapes that are not economically valuable for production (Pressey, 1994; Balmford and Whitten, 2003). As a consequence of bias towards landscapes of low productivity, areas of high priority for nature conservation tend to be located on unprotected private lands (Knight, 1999). Protecting privately-owned land for biodiversity involves many challenges, including the acceptance of additional conservation efforts by the landowners.

Fragmentation can intensify the effects of pure habitat loss producing potential population decline greater than expected solely on the basis of remaining area (Andrén, 1994; Bender et al., 1998) because of the loss of landscape connectivity. Connectivity has been linked to various ecological processes: it affects species colonisation and dispersal success (van Langevelde, 2000; Moilanen and Nieminen, 2002; Walters, 2007), extinction risk (Franken and Hik, 2004), population density (Fahrig and Paloheimo, 1988; Hanski et al., 1994) and population growth rate (Fahrig and Merriam, 1985). Connectivity can thus be considered to be a desirable quality of protected-area networks. Connectivity can only be understood in the light of two components. Structural connectivity refers to the degree to which some landscape elements are contiguous or physically linked to one another (With et al., 1997; Tischendorf and Fahrig, 2000), whereas functional connectivity recognises the behavioural responses of organisms to the physical structure of the landscape (Taylor et al., 1993; Bélisle, 2005). Thus, landscape connectivity depends not only on the amount and patterning of habitat, but also on the habitat affinities and dispersal abilities of species.

Modern forestry has severely altered the overall dynamics and structure of Fennoscandian forest ecosystems, although the total area of forest land has not changed (Östlund et al., 1997). Production forests are younger, more even-aged, more homogeneous, have less dead wood and deciduous trees than natural forests (Esseen et al., 1997). As a result of intensive forest management, many forest-dwelling species have become threatened. According to the

---

* Corresponding author. Tel.: +358 40 5419299; fax: +358 14 260 2321.
E-mail address: anne.laita@bytl.jyu.fi (A. Laita).

0006-3207/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biocon.2010.02.029

latest assessment, 38% of the red-listed species in Finland are forest species, and about 8% of all assessed forest species are threatened (Rassi et al., 2001). A large number of forest species has adapted to living under conditions that are not met in intensively managed forests.

Also in Fennoscandia, the majority of protected areas are located at high elevations and high latitudes, or in remote landscapes with low economic value (Nilsson and Gotmark, 1992; Virkkala, 1996; Stokland, 1997). Protected-area networks in more productive regions of Fennoscandia appear to be inadequate, and there is an imminent need to supplement the networks, particularly for forests (Angelstam and Andersson, 2001). In reaction to the notion that protected areas are not sufficient for forest biodiversity conservation, new forestry practices have been implemented in the managed forests of Fennoscandia. Woodland key habitats (WKHs) represent a new means in ecologically more sustainable forestry. WKHs are defined as habitats in which red-listed species are likely to occur (but not necessarily observed), and they are considered as sites harboring forest biodiversity (Hansson, 2001). WKHs are situated in managed forests, but still have many characteristics of natural stands. The idea in supplementing the network with WKHs is that although WKHs are usually rather small in size, they are putative biodiversity hotspots. Thus, they may represent a cost-efficient tool in safeguarding the forest biodiversity.

The Forest Act of Finland (1996/1093) defines a total of 13 distinct habitat types. The Forest Act obliges the forest owners to maintain the special features of these "habitats of particular significance", and allows only selective felling to be practiced on the sites. Alongside with traditional reserves, WKH-based conservation, as implemented in Finland, ensures the permanent protection of WKH sites. If not designated as WKH, the sites would be managed as typical commercial forests (thinned and clear-cut according to the prevailing recommendations). In the year 2007, delimited WKH sites covered ca. 85,000 ha (0.6% of the total forest land) of the private forests in Finland (Anonymous, 2008). In this study, the term WKH refers to the WKHs enjoying legal protection, although the concept of woodland key habitats in broad terms covers also other forest habitats besides the "Forest Act habitats". The other valuable habitats are preserved voluntarily by good forestry practices (Meriluoto and Soininen, 1998).

WKH-based conservation has received a lot of criticism, largely directed towards the small size of the WKHs and the associated aversive ecological consequences (e.g. Hansson, 2001) (covered in more detail in Section 4). The area of a delimited WKH is on average 0.62 ha (median 0.36 ha) (Yrjönen, 2004). It has also been argued that the WKHs form such a sparse network in a forest landscape that they cannot be important in biodiversity protection in the long term (e.g. Hanski, 2006; Pykälä, 2007). The too-small-too-isolated critique neglects the fact that WKHs are embedded in the landscapes together with the protected-area network. Protected areas and WKHs may have different functions in species conservation, and they should be viewed as conservation means that complement each other. Thus, the neteffect of WKHs on biodiversity hinges on their ability to support the network, and the contribution of WKHs to biodiversity can only be assessed if evaluated as integrated components in the landscapes. In the past decade, a lot of effort has been put to compare WKHs and more traditional protected areas (larger set-asides). We think it is time to evaluate the combined effects and efficiency of both of the set-aside types.

Graphs are abstractions of landscapes, where patches are represented by nodes connected by links. Links stands for functional inter-patch connections, and in the landscape-ecological context they usually represent the dispersal potential or the number of dispersing individuals between patches. Graph theory and its terminology in the field of landscape ecology have been reviewed by Bunn et al. (2000), Urban and Keitt (2001), as well as Urban and others (2009). A graph-theoretic approach makes it possible to analyse how patches are linked at various distance thresholds (i.e. dispersal abilities), which makes it possible to address the connectivity of the landscape from the species' perspective.

In this study we investigated the role of WKHs as an integral part of the traditional reserve network. We used a graph-theoretic approach to evaluate how and to what degree the WKHs support the reserve network by not only augmenting the protected habitat area, but most importantly, by creating functional connections among reserves. We tracked the amount of habitat that is functionally connected from the species' perspective to allow for spatially dependent ecological processes. We have used a habitat-based approach and not tried to account for any specific ecological processes (other than the dispersal potential among reserve patches). We think that this landscape-ecological approach may provide an insight into the value of WKHs not previously considered. We like to emphasise, however, that this is not a sufficient evaluation of the real ecological value of WKHs, as this is a product of many factors besides habitat connectivity (e.g. the habitat quality in WKHs, the effects of the surrounding matrix) (discussed later).

We wanted to shed light on the following questions: (1) How much do WKHs contribute to the connectivity of the network? (2) Are WKHs an efficient way to supplement the reserve network? Here we consider WKH-based conservation to be efficient if WKHs contribute to network connectivity more than expected by their contribution to habitat area. (3) What is the role of individual WKH patches in the network? The contribution of WKHs in the protection of rare and red-listed species obviously varies among species according to their habitat affinities and dispersal abilities. Thus we address question (4) What kinds of species benefit the most from the WKHs (in terms of habitat specialisation and dispersal ability)? We analysed the resulting network separately for different habitat types. In addition, to encompass a wide spectrum of species dispersal abilities, the networks were scrutinised using several threshold distances among patches varying between 200 m and 25 km.

It is expected that the value of WKHs may depend on the level of habitat availability in the landscape. For that reason, we compared the contribution and the role of WKHs in three areas which differed in terms of their reserve coverage and density of WKHs.

2. Materials and methods

2.1. Study area

This study was conducted in Central Finland, which extends some 240 km in the south–north direction (Fig. 1), and where forest characteristics vary in different parts of the region. In the region, forested land (including productive forest land, scrubland and barren land) covers 85.5% of the total land area, and 96.6% of this forest land is used for forestry (Korhonen et al., 2007). Pine-dominated (*Pinus sylvestris* L.) forests cover 59%, spruce-dominated (*Picea abies* (L.) Karst.) 31% and deciduous-dominated (mainly birch, *Betula pendula* Roth and *B. pubescens* Ehrh.) 9% of the productive forest land.

Due to the strict privacy protection policy of forest owners in Finland, there is no general access even for researchers to the data containing the exact locations and delimitations of WKHs. For this reason, we were obliged to restrict our analysis to three separate study areas (Fig. 1) for which data were obtainable by special negotiations. Our study areas are not a random sample of the region, but represent the overall variation in landscape structures and WKH densities across the region. Each study area covers approximately 500 km^2 (Table 1).

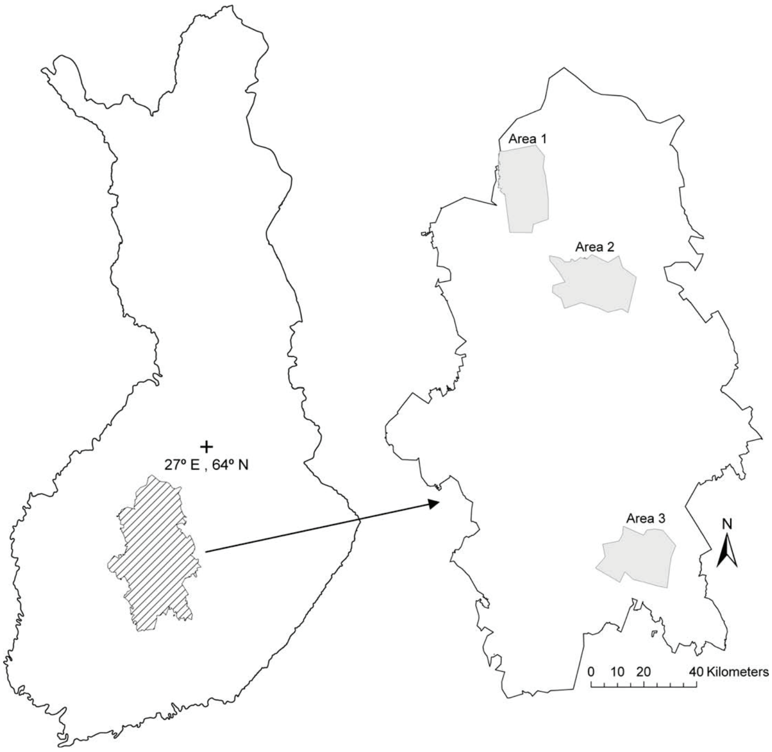


Fig. 1. The three study areas located in Central Finland.

Table 1
Areal information for the three study areas.

Measure	Area 1	Area 2	Area 3
Total area (km^2)	507	517	465
Total land area (km^2)	454	487	374
Area of forests, scrubland and barren land (km^2)	431	483	347
Total area of reserves (ha)	5749	2244	4030
Coverage of reserves (%)[a]	13.3	4.6	11.6
Number of WKHs	195	279	533
Total area of WKHs (ha)	165	215	310
Coverage of WKHs (%)[a]	0.38	0.45	0.89

[a] % of area of forests, scrubland and barren land.

Area 1 is located in the middle boreal vegetation zone and represents the low-productive watershed divide area of Suomenselkä, characterised by sparsely forested mires, xeric heath forests and rocky areas. In Area 1, forests cover 65.0%, mires and bogs 9.7% and other scrublands and barren land 18.3% of the total land area (based on the national CORINE database of Finland, spatial resolution 25 m).

Area 2 is situated in the transitional zone between the southern and middle boreal vegetation zones. This region is characterised by numerous lakes and small water bodies. Forests cover 76.2%, mires and bogs 4.9% and other scrublands and barren land 15.7% of the total land area.

Area 3 is located in the southern boreal vegetation zone. This southern area of Central Finland hosts many herb-rich forests and other fertile habitat types. The vegetation in this area is more varied than in other parts of the region (Uusitalo and Paakkolanvaara, 2007). Varied rock areas are also typical of this area. In Area 3 forests cover 64.6%, mires and bogs 4.5% and other scrublands and barren land 21.2% of the total land area.

The proportion of strictly protected areas (of the total area of forests and scrublands) in the whole of Finland is 8.9% (Southern Finland 2.2%; Northern Finland 15.6%). The share of strictly protected areas in Central Finland is 1.9%. When protected areas where cautious felling is allowed (WKHs, habitat types protected under the Nature Conservation Act and areas set aside by industry) are considered, the share of protected areas rises to 9.5% in the whole of Finland (Southern Finland 2.7%; Northern Finland 16.3%; Central Finland 2.5%). In our study areas, protected areas are more common than in Central Finland in general (Table 1) because all study areas host a national park, which increases their reserve coverage above the average. In Central Finland, WKHs on private lands cover 0.4% of the forestry land (compared to 0.6% in the whole country, 0.4% in Southern Finland and 0.8% in Northern Finland) (Anonymous, 2008). Coverages of WKHs in our study Areas 1 and 2 are on the average level, but Area 3 comprises more WKHs than average forest landscapes.

2.2. Landscape and habitat data

Data on the habitat types and exact locations of WKHs in the study areas were obtained from Metsähallitus (state-owned land), forest companies in the region, and from the Forest Centre in Central Finland (privately-owned forests). Fourteen landowners denied permission to data on their forest holdings, which excluded 68 WKH sites from our analyses. The excluded sites comprise 6.2% of the total number of WKHs in our three study areas, and all WKH-habitat types were represented in them. Because of this, our analyses will be conservative and the overall contribution of WKHs to the network connectivity for all habitat types is in reality somewhat greater than our results indicate.

The protected-area network included national parks, nature conservation program reserves, privately-owned protected areas, Natura 2000 areas, habitat types protected under the Nature Conservation Act, protected areas included in the regional plan, and sites protected by forest companies' own decisions. Hereafter we shall refer to all these other protected areas besides WKHs as 'reserves'. The habitat type information for the reserves was obtained from the GIS-based database administered by Metsähallitus, which contains detailed habitat type classifications for all patches in reserves. For some privately-owned reserves, habitat type information was obtained from the Regional Environment Centre and from the Regional Council of Central Finland.

Habitat type information was classified to form six habitat networks (Tables 2 and 3). All adjoining patches of the same habitat type were aggregated to present the network from the species' point of view. The reserves and WKHs were dealt with separately in the analyses, however. The networks formed a hierarchical organisation. Forest and peatlands were highest in the hierarchy, and the networks in the lower levels were subclasses of the more general networks. Networks also overlapped with each other. For example, forests and peatlands shared spruce-birch fens that foster species that are common to both habitat types. Herb-rich spruce mires were also included in herb-rich forest network as they host species typical of herb-rich forests in addition to their own characteristic species.

A WKH can comprise more than one habitat type. For example, the immediate surroundings of brooks can also have characteristics of herb-rich forests and herb-rich spruce mires. Usually the habitat subtypes are also registered in the database, and we treated a WKH as representative of all the habitat types recorded.

2.3. Network analyses

We calculated the overall network connectivity for all six habitat networks with a threshold distance ranging from 200 m to 25 km. The increment of threshold distance was calculated as $x_{h+1} = x_h \cdot e^{0.4}$, starting with the minimum distance threshold of 200 m. This yielded a (rounded) series of distance thresholds: 200 m, 300 m, 400 m, 700 m, 1 km, 1.5 km, 2 km, 3 km, 5 km, 7 km, 10 km, 15 km and 25 km. We wanted to make the increments larger with increasing threshold distance. For poor dispersers, an increment of 100 m may have a great influence on functional connectivity, but for good dispersers only increments of kilometres in distance will matter.

For every habitat type two networks were constructed: one without WKHs and the other with WKHs. By comparing the two networks, we inferred the contribution of WKHs to the network connectivity at different threshold distances. We calculated the %-increase in connectivity that was produced by the addition of WKHs to the network. We also calculated the proportion of the overall network connectivity that was contributed by the WKHs.

Networks were analysed as undirected graphs. Link weights were determined based on Euclidean edge-to-edge distances between patches. The network interpretations were based on the Integral Index of Connectivity (IIC) (Pascual-Hortal and Saura, 2006, 2008), which ranges from 0 to 1 and is computed as:

$$\mathrm{IIC} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_i \cdot a_j}{1+l_{ij}}}{A_L^2}$$

Table 2
Habitat networks analysed in our study. Habitat network refers to our reclassification of habitat types in reserves and WKH types outside reserves to provide six networks. The networks are to some extent hierarchical as, for example, "Forests" include habitat types and WKHs of "Heath forests".

Habitat network	Habitat types in reserves	WKH types
Peatlands	Pine mires, fens, eutrophic fens and spruce-birch fens	• Sparsely forested mires • Eutrophic fens • The immediate surroundings of small ponds (if not defined as flooded meadow)
Forests	Heath forests, herb-rich forests, spruce mires, wooded swamps and forests on rocky terrain	• Immediate surroundings of brooks, rivulets and springs (in forest area) • Fertile patches of herb-rich forests • Herb-rich spruce mires • Heathland forest islets in undrained peatland • Steep bluffs and the underlying forest stands • Gorges and ravines • Sandy soils, exposed bedrocks and boulder fields
Heath forests	Heath forests (barren, xeric, sub-xeric, mesic, herb-rich), herb-rich forests and spruce mires (excluding spruce-birch fens)	• Immediate surroundings of brooks and rivulets • Fertile patches of herb-rich forests • Herb-rich spruce mires • Heathland forest islets in undrained peatlands • Steep bluffs and the underlying forest stands
Herb-rich forests	Herb-rich forests, herb-rich heath forests and herb-rich spruce mires	• Fertile patches of herb-rich forests • Herb-rich spruce mires
Spruce-mires	Spruce mires and spruce-birch fens	• Herb-rich spruce mires
Herb-rich spruce mires	Herb-rich spruce mires	• Herb-rich spruce mires

Table 3
Number of patches and total area (ha) of different habitat types in WKHs and reserves for the three study areas.

Biotope	Area 1		Area 2		Area 3	
	WKH sites	Reserves	WKH sites	Reserves	WKH sites	Reserves
Peatlands						
N	53	123	89	82	113	137
Area total	46.2	2738.3	96.2	756.1	88.4	1461.3
Forests						
N	123	309	162	165	355	301
Area total	117.3	2664.0	127.8	1400.6	224.0	2516.7
Heath forests						
N	62	317	74	156	242	286
Area total	69.7	2551.5	78.2	1367.0	155.9	2478.2
Herb-rich forests						
N	25	17	17	3	112	116
Area total	28.3	50.2	17.2	8.0	67.4	312.9
Spruce mires						
N	24	134	10	59	43	163
Area total	27.3	255.8	14.2	180.7	28.9	174.8
Herb-rich spruce mires						
N	24	12	10	0	43	38
Area total	27.3	34.2	14.2	–	28.9	27.0

where n is the total number of patches, a_i and a_j are the sizes of patches i and j, l_{ij} is the number of links in the shortest path between patches i and j, and A_L is the total area of forested land (forests, scrubland and barren land) in the landscape. IIC reaches unity when all of the forested land is occupied by the given habitat type.

IIC is based on a binary connection model, and it treats two patches as connected if a link weight is below a threshold distance and otherwise as unconnected. IIC appreciates the habitat availability concept by integrating the habitat amount and connections between patches in a single measure. This approach recognises that connectivity also happens within a habitat patch, not only between patches. Area-informed indices avoid the ecologically counterintuitive outcome that network connectivity increases with the number of patches, irrespective of the total patch area.

In addition to the overall network evaluation, we calculated patch importance and centrality measures for individual patches. With the patch-level measures we wanted to analyse the role of WKHs in the network based on comparisons between WKH and reserve patches. Patch importance and network centrality measures were calculated at the distance thresholds of 500 m, 1 km, 2.5 km, 5 km, 7.5 km and 10 km. Patch importance values were calculated with the node removal analysis: each patch is systematically removed from the network and the reconstructed network is compared with the original network including the patch. The connectivity loss (per cent) measures the contribution of the patch to the network connectivity. Patch importance values were calculated based on IIC connectivity. We also calculated the per area based importance values by dividing the patch importance value by the patch area. Inferences were based on averaged patch importance values over all the threshold distances.

The network centrality of nodes was evaluated with two measures: degree and betweenness centrality. The degree of a patch equals the number of its direct neighbors. It is a good and simple measure of determining how well a patch is connected to other patches at a local scale (Estrada and Bodin, 2008).

The betweenness centrality of a patch is the proportion of the shortest paths (based on topological distance) between all pairs of patches which run along the patch. For patch x_i the betweenness centrality is calculated as

$$C_B(x_i) = \sum_{j<k}^{n}\sum^{n} b_{jk}(x_i)$$

where $b_{jk}(x_i)$ is the proportion of the shortest paths linking x_j and x_k that contain x_i and $i \neq j \neq k$ (according to Freeman et al. (1991)). The contribution of a patch to the large-scale connectivity of the landscape increases with the betweenness centrality measure (Estrada and Bodin, 2008). Betweenness centrality can be used to point out the patches that are located in critical positions relative to other patches from the point of view of a functional network, and has been used to identify stepping stone patches in a patch network. If these critical patches (cut-nodes in graph terminology) were removed, the network would risk being dissected into isolated groups of patches.

We confined our analyses to top-ranking patches in betweenness centrality, because low to intermediate centrality scores do not necessarily hold any meaningful information about the role of the patch in the network (Estrada and Bodin, 2008). We calculated the average for betweenness centrality over all the threshold distances. The patches with an average value >0.01 were designated as top-patches. To check if WKHs contribute to betweenness centrality more than expected by chance, we calculated the probability of finding k WKH patches among n top-patches based on binomial distribution with the following formula:

$$P(X = k) = (n; k)p^k(1 - p)^{n-k}$$

where p is the proportion of WKH patches in a network.

Graphs were analysed with Conefor Sensinode v.2.2 (Saura and Torné, 2009) and Pajek v.1.23 (de Nooy et al., 2005).

3. Results

The patch sizes in reserves were generally consistently larger than those in WKHs (Appendix A). The differences were not so marked, however, for fertile and rare habitat types (e.g. herb-rich forests, spruce mire, herb-rich spruce mires). The difference was even reversed for spruce mires in Area 2. In Area 2 herb-rich spruce mires only occurred in WKHs.

The influence of threshold distance on connectivity was evident in all the results. Networks experienced radical changes with increasing threshold distance; they were mainly composed of isolated patches at small threshold distances, whereas at larger distances they mainly occurred as one component with multiple pathways between patches (Fig. 2).

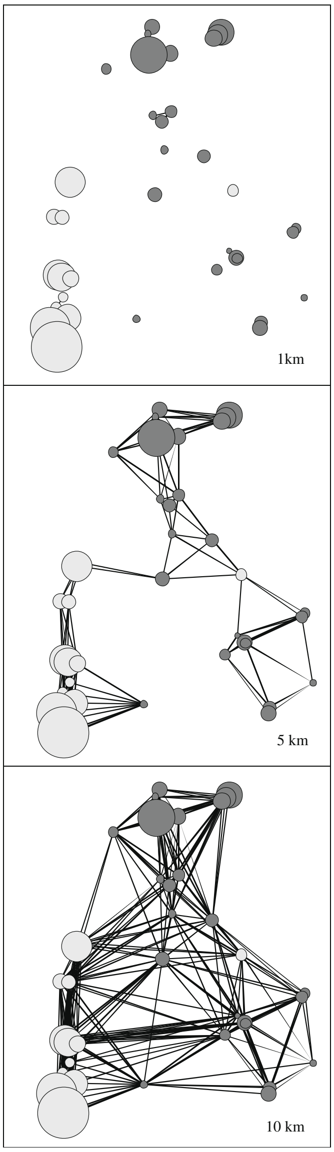


Fig. 2. The herb-rich forest network of Area 1. WKHs (in pale gray) and other reserves (in dark grey) are shown as spheres proportional to their patch sizes. This is a complete graph presentation of the habitat network, where the connections between patches are thresholded at link weights of 1 km, 5 km and 10 km.

For all habitat types, the connectivity of the network in Area 2 was remarkably lower than in other areas (Fig. 3.) Area 3 exceeded other areas in connectivity for forest habitats, whereas Area 1 had the highest connectivity for peatlands and spruce mires. The enhancement in connectivity from WKHs varied according to the threshold distance and among areas as well as habitat types (Fig. 3). This becomes evident when looking at the percentage increase in connectivity with increasing threshold distance calculated for all habitat types (Fig. 4). WKHs benefited mostly dispersal-limited species in networks of common habitat types (peatlands, forests, heath forests), for which the connectivity increase brought about by WKHs peaked at about threshold distances ranging from ca. 2 to 5 km. In networks of rare habitat types (e.g. herb-rich forests, spruce mires, herb-rich spruce mires), the relative contribution of WKHs to connectivity peaked at much larger threshold distances. But in sparse networks of rare habitat types, the level of the relative connectivity increase was much higher than in dense networks of common habitat types across all threshold distances (except the very smallest ones). This sparse-dense difference was also evident when different areas were compared with each other. Area 2, with small reserve coverage, gained generally more connectivity from WKHs than the other areas. For example, for the herb-rich forest network in Area 2, WKHs increased the habitat area by over 200%, leading to a connectivity increase of over 600%. For more dense networks, the connectivity increases were not so pronounced. For example, in networks of forests the areal increases of ca. 4–9% yielded connectivity increases of ca. 20–50%.

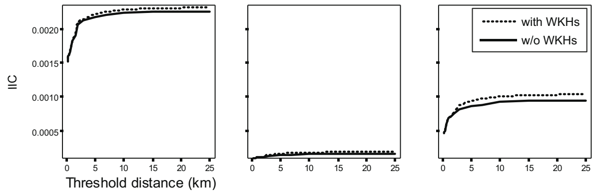
The efficiency of WKH-based conservation was highly dependent on the threshold distance. The proportion of the connectivity contributed by WKHs was smaller than their areal proportion (i.e., the proportion of WKHs of total network area) at small threshold distances, but with greater threshold distances the relative contribution exceeded the areal proportion in all cases (Fig. 5; see Appendix B for result summary). The threshold distance at which the relative connectivity contribution exceeded the areal proportion was network-specific, ranging between 1 and 10 km. For example, in the herb-rich forest networks of Areas 1 and 2, efficiency was achieved at the threshold distance of 5 km, whereas the networks of common habitat types reached the efficiency at a smaller threshold distance, for example heath forests at 1–2 km.

At the level of individual patches, the reserves exceeded the WKHs in patch importance values calculated per area unit (Table 4) (area-informed IIC-patch importance values are summarised in Appendix C). The median values for patch importance were in every case larger in reserves than in WKHs. Reserve patches are thus in general more valuable than WKHs, not only because of their larger size, but also due to their favourable configuration. Some small-sized stepping-stone WKHs had high patch importance values per area unit, which in some cases raised the average patch importance values of WKHs to a higher level than in the reserves. For example, the herb-rich forest WKH patches in Areas 2 and 3 contributed, on average, more to network connectivity per area unit than their reserve counterparts. In networks of rare habitat types, the importance of one area unit was higher than in the networks of common habitat types. For example, one hectare patch in the herb-rich spruce mire networks was approximately worth 3% of the overall network connectivity, whereas in the forest network, one hectare patch was worth 0.05–0.1%. Any single WKH patch in a network of rare habitat type may have a notable value, contributing several percent to the whole network connectivity.

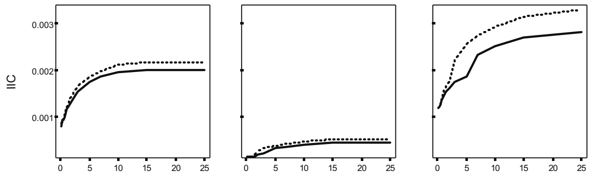
Reserve patches had on average more neighbors than WKHs at all threshold distances, and thus were better connected at the local scale. This applied, for example, to herb-rich forests in Area 3 (Fig. 6). Only in the herb-rich forest network in Area 2 did WKHs have on average a higher degree of scores than reserves across all scales. The degree centrality measures are given in Appendix D.

In sparse networks, proportionally more patches scored high in betweenness centrality than in dense ones. Individual patches in

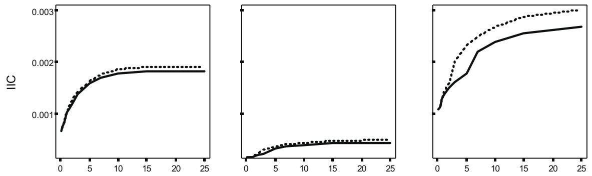
(a) Peatlands



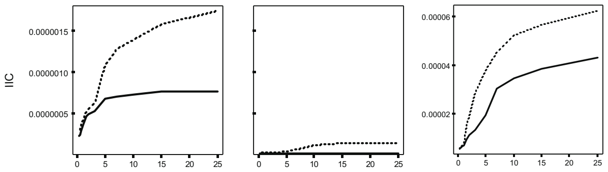
(b) Forests



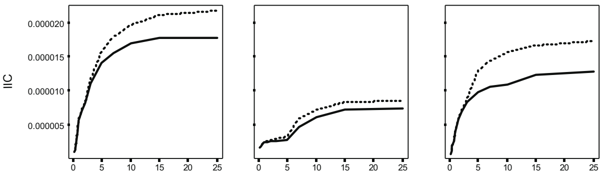
(c) Heath forests



(d) Herb-rich forests



(e) Spruce mires



(f) Herb-rich spruce mires

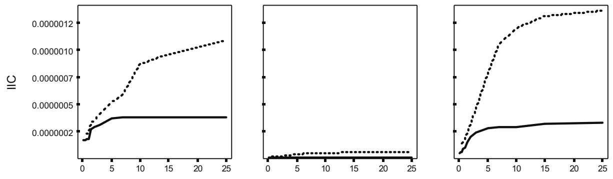


Fig. 3. Overall network connectivity based on the Integral Index of Connectivity (IIC) for different habitat networks as a function of threshold distance (i.e. dispersal ability). Two networks are formed; one without WKHs and another with WKHs.

dense networks did not play such a critical role in betweenness due to numerous alternative connections. The proportion of top-patches (of the total number of patches in an area) ranged from 2% in the heath forest network (network size 656 patches) of Area

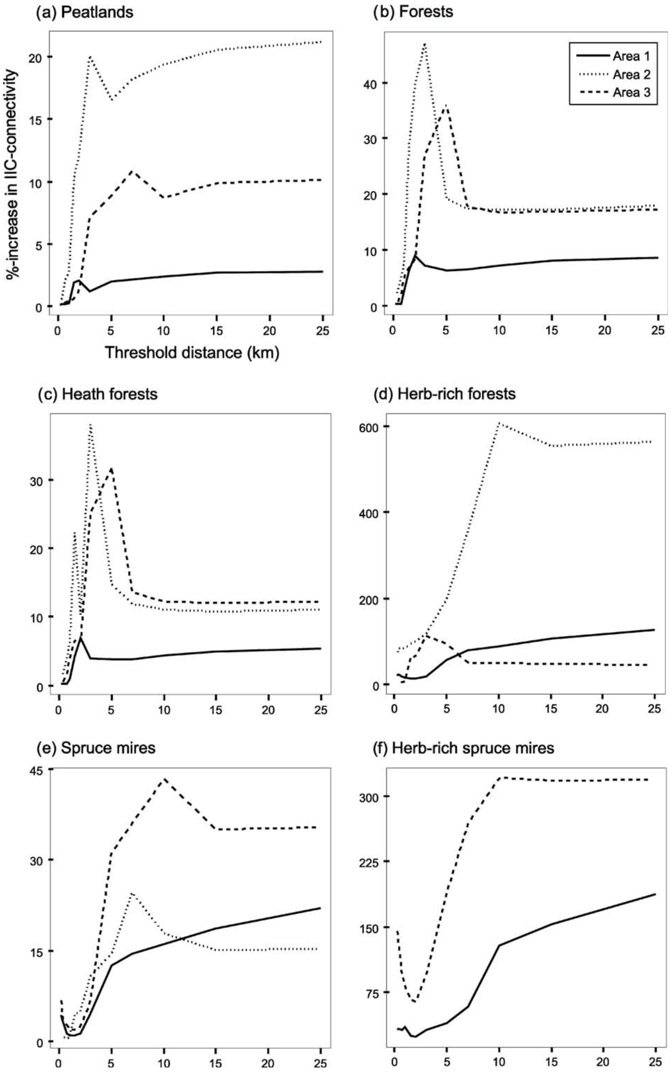


Fig. 4. Increase in IIC connectivity (per cent) brought about by the additions of WKHs to the different reserve networks as a function of threshold distance.

3–55% in the herb-rich forest network (network size 20 patches) of Area 2. The median proportion for top-patches for all area-habitat type network combinations was 11.6%.

In general, top-patches were distributed among reserves and WKHs in direct proportion to their numbers. However, in the forest, spruce mire and herb-rich spruce mire networks of Area 3, there were more WKHs among top-patches than expected by chance ($p = 0.022$, $p = 0.014$ and $p = 0.019$, respectively).

4. Discussion

When analysed from the point of view of a functional network, WKHs seem to be valuable. WKHs did not just contribute to the reserve network in terms of area, but they also created new interpatch connections. To ensure the availability of habitats that occur naturally scattered in a landscape, WKHs may be a more efficient and straightforward form of conservation than large reserves. We

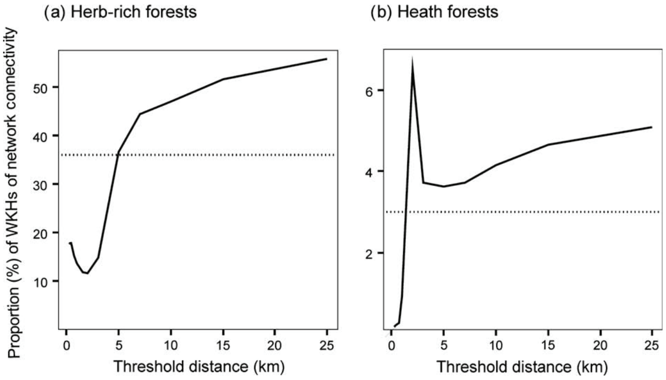


Fig. 5. The proportion (%) of WKHs of overall network connectivity as a function of threshold distance for the herb-rich forest and heath forest network of Area 1. The proportion (%) of WKHs of the total network area is shown with a dashed line. WKHs are an efficient way to increase landscape connectivity when their proportional connectivity contribution exceeds their areal proportion in the network.

Table 4
Patch importance values per area unit (ha) (based on the Integral Index of Connectivity, IIC) compared between WKHs and reserves for three study areas. Results were WKHs have higher average patch importance values per area unit than reserves are in bold.

	Area 1		Area 2		Area 3	
	WKHs	Reserves	WKHs	Reserves	WKHs	Reserves
Peatlands						
Median	0.027	0.054	0.107	0.109	0.049	0.083
Mean ± SD	0.031 ± 0.01	0.053 ± 0.03	0.118 ± 0.05	0.127 ± 0.06	0.053 ± 0.02	0.086 ± 0.04
Forests						
Median	0.036	0.067	0.073	0.081	0.035	0.068
Mean ± SD	**0.101 ± 0.01**	**0.063 ± 0.47**	**0.107 ± 0.18**	**0.101 ± 0.05**	0.046 ± 0.06	0.064 ± 0.02
Heath forests						
Median	0.032	0.073	0.084	0.084	0.038	0.071
Mean ± SD	0.045 ± 0.09	0.068 ± 0.01	**0.898 ± 6.83**	**0.103 ± 0.05**	0.048 ± 0.05	0.068 ± 0.02
Herb-rich forests						
Median	1.174	2.313	4.472	5.320	0.355	0.389
Mean ± SD	1.416 ± 0.78	2.319 ± 0.53	**4.576 ± 1.76**	**2.319 ± 1.90**	**0.424 ± 0.42**	**0.406 ± 0.13**
Spruce mires						
Median	0.303	0.714	0.525	1.069	0.445	1.114
Mean ± SD	0.369 ± 0.27	0.750 ± 0.34	0.625 ± 0.21	0.919 ± 0.53	0.449 ± 0.12	1.088 ± 0.46
Herb-rich spruce mires						
Median	1.440	2.989	10.532	–	2.355	3.705
Mean ± SD	**3.214 ± 5.20**	**2.811 ± 0.61**	10.368 ± 2.24	–	3.340 ± 4.14	3.773 ± 1.62

suggest that WKHs provide a means to supplement the nature reserve network in rare and scattered habitat types. This was evident particularly for herb-rich forests and herb-rich spruce mires. Large continuous forest reserves do not provide area or connectivity for these habitat types, and WKHs are definitely needed both to increase habitat availability and to promote the dispersal of species specialised in such habitat types.

The efficiency of the WKH-based conservation depended greatly on the dispersal abilities of species. For species with weak dispersal capabilities, WKHs are not likely to be an optimal way of providing habitat, but larger set-asides would be a more efficient solution. For species with good dispersal capabilities, WKH-based conservation seems to be efficient when analysed from the network perspective. For these species the WKHs are also bridging habitats in reserves that would otherwise remain unconnected, and thus increase habitat availability over their own areal proportion. There is no specific limit for the dispersal capacity above which the WKH would prove to be efficient; instead, the outcome is dependent on the characteristics of the network. For species requiring more rare and scattered habitat types, a greater dispersal capacity is required in order for them to benefit from the contribution of WKHs to connectivity than for species requiring more common habitat types.

The result of our study, i.e., that the networks were in general considerably more connected with than without WKHs, points to the value of WKHs at the level of the whole network, not at the level of individual WKHs. Reserve patches were generally larger in size and more strategically located than WKHs. It is the combined effect of all the WKHs together with reserves that makes them valuable in the network. The relatively high density of WKHs is the most likely reason for their value in the network as a whole. However, WKHs among the top-patches in between centrality indicate that some individual WKHs are also important in sustaining the large-scale connectivity of the networks. Without WKHs,

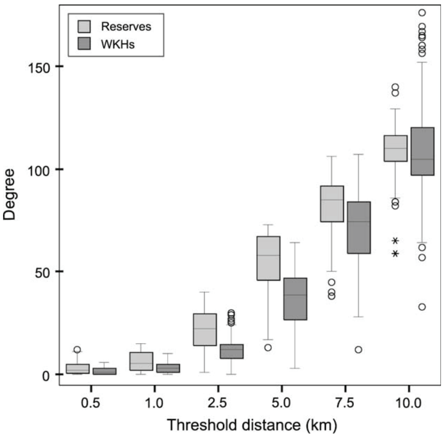


Fig. 6. Comparisons of degree centrality measures between WKHs and reserves for the herb-rich forest patches in Area 3 at threshold distances ranging from 0.5 km to 10 km.

dispersing individuals would experience the network as more fragmented.

The coverage of the reserve network in Area 2 is closest to the typical level in Southern Finland, although somewhat above the average. The share of WKHs in Area 2 represents the average level in Central Finland and is ca. 0.1% lower than in Southern Finland as a whole. In our study, Area 2 exemplified an area with a sparse reserve network, and in this area the connectivity of the reserve network was greatly enhanced by WKHs. This reflects the potentially important role of WKHs as part of the reserve network in Southern Finland. In regions with substantially greater reserve coverage, the WKH-based conservation most likely serves species that are highly dispersal-limited. The results from Area 3 indicate that the high density of WKHs increases the value of WKH-based conservation. This is most evident when comparing the results from Area 1 and Area 3 which cover approximately the same reserve area while WKHs are more numerous in Area 3.

The target species of WKHs are red-listed species with small areal demands, mainly bryophytes, invertebrates (though very poorly studied), lichens, polypores, and vascular plants. Empirical evidence suggests that WKHs may be biodiversity hotspots for epiphytic lichens (e.g. Johansson and Gustafsson, 2001; Pykälä, 2004) and rare bryophytes (Perhans et al., 2007; but see Gustafsson et al., 2004), but not necessarily for polypore fungi (Sippola et al., 2005). Dispersal has been considered a limiting factor for many threatened forest species, which call for spatial considerations in their protection (see Edman et al. (2004a) and Penttila et al. (2006) for polypores; Sillett et al. (2000) and Johansson and Ehrlén (2003) for epiphytic lichens).

The dispersal capacities of threatened forest species are poorly known, and thus we cannot draw firm conclusions about how large a proportion of the red-listed species are good enough dispersers to gain benefit from the enhanced connectivity provided by the WKHs. Our range of threshold distances (from 200 m to 25 km) is based on the dispersal estimates for threatened forest beetles, as they represent the extremes of reported dispersal abilities for threatened forest biota. The maximum detected dispersal distance of a hollow-tree specialist *Osmoderma eremita* was found to be 190 m (Ranius and Hedin, 2001), whereas Jonsson (2003) recorded a median dispersal distance of 12 km and a maximum distance of 28.7 km for the saproxylic beetle *Oplocephala haemorrhoidalis* in a flight-mill experiment. When more information on species dispersal distances becomes available, our results can be interpreted a posteriori for any given threatened species with an estimate of its dispersal ability. It is worth noting that a functional reserve network cannot be based on extreme dispersal events. For example, although the individual airborne spores of polyporous fungi can travel very long distances, the successful colonisation (requiring large number of spores) of polypore species has been suggested to be dispersal-limited (Edman et al., 2004a,b). Thus, it is important that the dispersal ability estimates used for interpretations are ecologically realistic and not overly optimistic.

To be functionally connected, reserve networks (even with WKHs) place great demands on the dispersal ability of organisms. If a species' maximum dispersal distance is 3 km, all habitat type networks are unconnected (based on IIC). Networks of peatlands, forests and heath forests are very close to being connected with this threshold distance, but individuals would still perceive the other networks disconnected. It seems that the sparser networks are too fragmented for dispersal-limited forest species. The WKH patches considerably increased the connectivity of the sparse networks, but the increase in connectivity benefits most the species with a good dispersal ability (up to 5 km). It is kilometres that matter in these networks, not hundreds of metres. For poor dispersers, the networks with or without WKHs consist mainly of isolated patches.

The binary approach to a landscape consisting of a matrix surrounding the suitable ''islands'' of habitat patches is appropriate for the species that perceive the matrix as predominantly hostile. Many red-listed species have such specific habitat requirements (for example, those requiring a shady microclimate or a high density of dead-wood) that they cannot easily be fulfilled in production forests. The degree to which the reserves are functionally linked by dispersal becomes increasingly important for those specialised forest species. By contrast, the species for which the matrix also provides habitats and resources, perceive the landscape as a more gradually varying entity of different resource densities. For such species, our analysis may severely underestimate habitat availability and landscape connectivity.

In our study, the distances between patches were calculated as Euclidean distances. As the target species of WKHs mainly disperse by airborne spores and seeds, a geographical distance is a good approximation of the inter-patch distance experienced by an organism. For those species that use insect vectors or active movement (shaped by the matrix) for their dispersal, the straight-line distance may underestimate the effective distance between patches.

We analysed the value of WKHs only from the habitat connectivity perspective. There are also other things that need to be considered to determine the ultimate value of WKHs in the conservation of threatened species. There are many qualitative problems in setting aside small forest parcels embedded in a managed forest matrix. Many of the problems are directly related to the small size of the WKHs. Our study was conducted on the assumption that WKHs contribute to connectivity worth their habitat area. Ecological processes are, however, usually weighted down in small patches surrounded by an intensively managed matrix, so the habitat area does not necessarily equal the ecological value. The smaller the patch is, the greater the influence external factors are likely to have on the microclimatic conditions (Saunders et al., 1991). According to a study by Aune et al. (2005) most of the small-sized WKHs totally lack core area. The areas of WKHs based on official delimitations may be too optimistic, especially from the point of view of species that are sensitive to microclimatic conditions.

Small WKH sites can host only small populations, and thus the populations in WKHs are prone to extinctions (e.g. Hansson, 2001). Pykälä (2004) found in his monitoring study on epiphytic macrolichens that although concentrated in WKHs, their small population sizes predisposed them to local extinctions in less than 10 years. The long term occurrence of species in WKHs may thus depend on repeated colonisation to compensate for local extinctions. Colonisation rate has been shown to depend on the isolation of a patch (e.g. Verboom et al., 1991). The immigration of individuals to a small population may save that population from extinction (called 'the rescue effect') (Brown and Kodric-Brown, 1977). The rescue effect, and thus also extinction risk, is also related to the isolation of patches as the possible immigration diminishes with increasing distance from the other occupied patches. The configuration of suitable patches may thus be a critical issue for the viability of populations residing in WKHs. In our study, we evaluated the connectivity contributions of WKHs from the point of view of the reserve network, but the traditional reserve network is likewise a prerequisite for successful WKH-based biodiversity protection.

Finnish legislation allows cautious selective logging in WKHs as long as site features are not destroyed or altered. This may cause reductions in ecological values, a decrease in the amount of dead wood in long term, for example. Selective logging can also change the microclimatic conditions and make them unsuitable for most sensitive species. Pykälä (2004) found that logging in WKHs was the main cause of extinctions in epiphytic macrolichens. To sustain population persistence and to augment the existing reserve network, it is important that WKHs maintain their habitat qualities. It may be necessary to refrain from any logging in WKHs. Similarly, a buffer zone around WKHs would give protection against the detrimental effects that a surrounding matrix exerts on WKHs. For example, 30-m buffer zones have been found adequate to maintain the original species composition in riparian WKH sites (Selonen and Kotiaho, 2006).

It has been suggested that there is a threshold value of habitat availability below which the effect of habitat patterns on population persistence may become evident (Andrén, 1994). For example, Andrén (1994) reported an empirical threshold value of 10–30% of habitat availability for birds and mammals. The threshold value is species- and landscape-context-specific and supposedly higher for species dispersing less well than birds and mammals (Mönkkönen and Reunanen, 1999). As the coverage of reserves in Finland apparently falls below any critical threshold, the configuration of reserves in the landscape is important for the persistence of many threatened forest species. Thus, WKHs have a potentially important role in providing connectivity for species for which production forests are unsuitable.

The importance of reserve configuration is scale dependent. For organisms that perceive the landscape at fine scales, landscape configuration is of little consequence, because populations are restricted to local habitat patches (Keitt et al., 1997). For these poor dispersers the importance of a reserve is thus largely determined by its size. Likewise, landscape configuration is likely to be of minor importance for species capable of traversing long distances across hostile landscapes. In contrast, species with a dispersal ability within a critical threshold range (i.e. the range where a small change in dispersal ability produces a great change in connectivity) experience the importance of spatial configuration the most (Keitt et al., 1997). WKHs shift the critical threshold range towards lower threshold distances, and species with a dispersal ability within this range shift are the ones that WKH-based conservation serves the best. As the critical threshold range is network-specific (for common habitat types occurring at lower threshold distances than for rare habitat types), the benefits associated with any given dispersal ability vary among different networks.

We suggest that WKH-based conservation would bring the greatest benefits in landscapes with a rather low habitat availability. Thus, setting aside WKHs would seem to be a more efficient and well-founded conservation tool in landscapes with intensive land use and forest management history. By contrast, we see little value in WKHs in landscapes where habitat availability is still at a high level for most species. However, in landscapes with a very low habitat availability and a highly fragmented habitat network, WKH-based conservation may not serve dispersal-limited species but only efficient dispersers instead. The benefits of WKH-based conservation can be evaluated in the light of the landscape context and the dispersal abilities of the species.

5. Conclusions

The smaller the land parcels we are setting aside, the more important the issue of functional connectivity becomes. New approaches are needed that deal explicitly with the spatial arrangement of the reserves in order to form ecologically sustainable and functional networks. To effectively maintain biodiversity through time, a reserve network needs to be designed and evaluated with explicit consideration to the reserve configuration. Woodland key habitats have been criticised as being too small and scattered in distribution. However, WKHs are not detached conservation elements in the forest landscape. We argue that they form a network with other reserves, and their real ecological value can only be understood as an integral part of the network.

Our main aim here was to evaluate the importance and role of WKHs within the reserve network. Our results show that, especially for species requiring fertile habitat types that occur naturally scattered in the forest landscape, WKH-based conservation can be very valuable. The value of setting aside WKHs is strongly linked with the dispersal ability of threatened species. It is evident in our analysis that the protection of large set-asides would be a more efficient conservation strategy for species with weak dispersal abilities. From the perspective of the functional connectivity in the protected-area network, the whole is clearly more than the sum of its constituent parts. Network connectivity is an emergent property that can only be understood at the level of the whole network.

Finally, even though we found that WKHs are important from the connectivity perspective, we note that the conservation of small-sized WKHs surrounded by an intensively managed forest matrix does also have its pitfalls. Most importantly, the site characteristics of WKHs must be safeguarded for them to be able to contribute to the connectivity of the reserve network. The deleterious effects of selective logging should be prevented in the first place. Buffer zones around WKHs would also moderate the edge effects from the surrounding matrix on WKH sites. On these conditions, WKHs seem to have the potential to contribute to network connectivity up to the level shown in our study.

Acknowledgements

This research was conducted as part of The Finnish Environmental Cluster Research Programme "Talousmetsät, pienialaiset suojelukohteet ja suojelualueet monimuotoisuuden suojelussa: Integroitu ekologinen vaikuttavuus ja kustannustehokkuus". The study was partially funded by a grant from Kone Foundation (to AL). We are also grateful for the funding provided by the Academy of Finland (Project #7115560, to MM). We thank the Forest and Park Service, the Regional Forestry Centre in Central Finland, the Regional Council, UPM Kymmene and the Regional

Environment Centre for providing data for our study. Special thanks to Seija Tiitinen-Salmela (The Forestry Centre in Central Finland) for contacting the landowners and to all landowners who let us to use information on their property in this study.

Appendix A

Mean (±SD) and median patch sizes of different habitat types in WKHs and reserves for three study areas.

Biotope	Area 1		Area 2		Area 3	
	WKH sites	Reserves	WKH sites	Reserves	WKH sites	Reserves
Peatlands						
Mean patch size ± SD	0.87 ± 0.75	22.26 ± 119.03	1.08 ± 1.68	9.22 ± 43.43	0.78 ± 1.27	10.67 ± 59.53
Md for patch size	0.59	1.53	0.63	1.82	0.43	0.95
Forests						
Mean patch size ± SD	0.95 ± 1.43	8.62 ± 31.69	0.79 ± 1.08	8.49 ± 27.43	0.63 ± 0.95	8.36 ± 38.50
Md for patch size	0.58	1.07	0.41	1.00	0.34	0.57
Heath forests						
Mean patch size ± SD	1.12 ± 1.50	8.05 ± 24.46	1.06 ± 1.21	8.76 ± 27.54	0.64 ± 0.86	8.66 ± 34.01
Md for patch size	0.68	1.07	0.63	1.05	0.37	0.62
Herb-rich forests						
Mean patch size ± SD	1.13 ± 1.40	2.95 ± 3.57	1.01 ± 0.80	2.67 ± 3.11	0.60 ± 0.94	2.70 ± 6.60
Md for patch size	0.77	1.18	0.81	0.95	0.33	0.77
Spruce mires						
Mean patch size ± SD	1.14 ± 1.43	1.91 ± 2.51	1.42 ± 0.77	3.06 ± 5.68	0.67 ± 1.15	1.07 ± 1.62
Md for patch size	0.75	1.10	1.42	1.31	0.38	0.60
Herb-rich spruce mires						
Mean patch size ± SD	1.14 ± 1.43	2.85 ± 3.33	1.42 ± 0.77	–	0.67 ± 1.15	0.71 ± 0.84
Md for patch size	0.75	1.23	1.42	–	0.38	0.47

Appendix B

The proportion (%) of WKHs of total network area and of overall network connectivity at different threshold distances. Results were WKH have higher network connectivity than their areal proportion are in bold.

	Area 1	Area 2	Area 3
Peatlands			
Area	1.66	11.29	5.71
Overall			
200 m	0.11	0.32	0.10
1 km	0.27	2.72	0.37
2 km	2.04	9.26	1.10
3 km	1.18	**16.75**	**6.64**
5 km	**1.96**	**15.37**	**8.16**
10 km	**2.30**	**16.23**	**8.04**
25 km	**2.67**	**17.49**	**9.19**
Forests			
Area	4.22	8.36	8.17
Overall			
200 m	0.24	1.78	0.16
1 km	2.51	7.22	4.97
2 km	**8.14**	**28.27**	7.41
3 km	**6.71**	**32.04**	**21.27**
5 km	**5.85**	**16.08**	**26.58**
10 km	**6.76**	**14.68**	**14.31**
25 km	**7.89**	**15.14**	**14.75**

(*continued on next page*)

Appendix B (*continued*)

	Area 1	Area 2	Area 3
Heath forests			
Area	2.66	**5.41**	5.92
Overall			
200 m	0.17	1.34	0.11
1 km	0.90	**6.13**	3.75
2 km	**6.50**	**9.37**	**5.87**
3 km	**3.73**	**27.60**	**19.98**
5 km	**3.73**	**12.86**	**24.18**
10 km	**4.16**	**9.94**	**10.84**
25 km	**5.08**	**9.92**	**10.92**
Herb-rich forests			
Area	36.00	68.22	17.72
Overall			
200 m	17.64	40.48	2.40
1 km	13.47	45.91	7.30
2 km	11.62	49.63	**39.16**
3 km	14.84	54.06	**53.11**
5 km	**36.54**	**66.28**	**48.54**
10 km	**47.00**	**85.83**	**33.46**
25 km	**55.75**	**84.95**	**31.10**
Spruce mires			
Area	9.64	7.26	14.20
Overall			
200 m	4.23	0.59	6.97
1 km	1.02	0.67	1.91
2 km	1.35	4.79	2.44
3 km	4.39	**9.81**	6.09
5 km	**11.11**	**12.63**	**23.76**
10 km	**13.92**	**15.21**	**30.23**
25 km	**18.02**	**13.22**	**26.13**
Herb-rich spruce mires			
Area	44.37	100	51.73
Overall			
200 m	24.04	–	**59.84**
1 km	25.42	–	44.45
2 km	18.90	–	38.71
3 km	23.52	–	48.63
5 km	27.83	–	**65.34**
10 km	**56.27**	–	**76.27**
25 km	**65.22**	–	**76.16**

Appendix C

Patch importance values (based on Integral Index of Connectivity, IIC) compared between WKHs and reserves for three study areas.

	Area 1	Area 2	Area 3			
	WKHs	Reserves	WKHs	Reserves	WKHs	Reserves
Peatlands						
Min–max	0.002–0.081	0.001–63.329	0.007–1.020	0.006–83.617	0.001–0.540	0.001–78.033
Median	0.018	0.084	0.071	0.176	0.022	0.076
Mean ± SD	0.025 ± 0.02	1.286 ± 7.05	0.118 ± 0.14	1.719 ± 9.34	0.042 ± 0.07	1.102 ± 6.84
Sum	1.326	158.23	10.479	140.945	4.753	150.979
Forests						
Min–max	0.000–0.361	0.001–29.442	0.002–0.878	0.002–44.115	0.001–0.302	0.009–39.958
Median	0.022	0.067	0.029	0.095	0.014	0.031
Mean ± SD	0.039 ± 0.059	0.619 ± 2.549	0.073 ± 0.116	1.026 ± 4.022	0.027 ± 0.041	0.587 ± 3.172
Sum	4.858	191.121	11.756	169.304	9.587	176.631

Appendix C (*continued*)

	Area 1	Area 2	Area 3			
	WKHs	Reserves	WKHs	Reserves	WKHs	Reserves
Heath forests						
Min–max	0.001–0.317	0.001–17.438	0.003–2.599	0.002–46.054	0.001–0.254	0.001–30.841
Median	0.022	0.070	0.055	0.100	0.016	0.038
Mean ± SD	0.042 ± 0.06	0.615 ± 2.06	0.147 ± 0.33	1.104 ± 4.28	0.030 ± 0.04	0.640 ± 2.89
Sum	2.584	194.966	10.849	172.302	7.211	182.933
Herb-rich forests						
Min–max	0.212–10.632	0.715–42.583	0.172–15.557	2.414–42.583	0.007–2.716	0.004–36.847
Median	0.873	1.900	4.174	5.070	0.116	0.310
Mean ± SD	1.682 ± 2.28	7.530 ± 10.51	5.377 ± 5.21	16.689 ± 22.46	0.223 ± 0.36	1.245 ± 3.74
Sum	42.049	128.016	91.407	50.067	24.991	144.408
Spruce mires						
Min–max	0.057–1.963	0.031–16.549	0.167–2.379	0.027–49.613	0.008–3.963	0.009–13.605
Median	0.252	0.775	0.637	0.933	0.159	0.622
Mean ± SD	0.345 ± 0.39	1.430 ± 1.20	0.890 ± 0.66	2.812 ± 6.74	0.322 ± 0.61	1.137 ± 1.70
Sum	8.286	191.671	8.895	165.914	13.832	185.366
Herb-rich spruce mires						
Min–max	0.255–14.892	0.954–43.492	1.379–29.260	–	0.033–27.031	0.030–3.194
Median	1.160	3.024	12.568	–	0.867	1.427
Mean ± SD	2.557 ± 3.26	9.064 ± 12.12	15.276 ± 9.24	–	2.122 ± 4.19	2.633 ± 3.19
Sum	61.36	108.765	152.764		91.225	100.040

Appendix D

Means (±SD) for degree centrality measure of different habitat types in WKHs and reserves for three study areas. Results where WKHs have higher average degree scores than reserves are in bold.

	Area 1	Area 2	Area 3			
	WKHs	Reserves	WKHs	Reserves	WKHs	Reserves
Peatlands						
500 m	1.1 ± 1.6	5.2 ± 5.3	1.2 ± 1.2	2.2 ± 2.2	2.2 ± 2.6	4.8 ± 3.4
1 km	2.7 ± 3.0	10.4 ± 7.4	3.2 ± 3.0	4.2 ± 3.2	5.5 ± 5.7	10.2 ± 5.2
2.5 km	10.2 ± 8.9	32.9 ± 16.8	11.7 ± 7.5	12.1 ± 7.0	17.8 ± 15.2	34.3 ± 13.0
5 km	30.7 ± 24.6	70.1 ± 24.4	**34.4 ± 15.0**	**33.3 ± 12.7**	44.8 ± 21.4	80.6 ± 19.9
7.5 km	54.7 ± 39.0	100.1 ± 25.3	58.6 ± 19.1	62.2 ± 20.3	81.1 ± 28.8	117.1 ± 19.8
10 km	83.1 ± 50.1	126.7 ± 21.4	82.7 ± 21.9	89.2 ± 25.1	125.5 ± 40.8	149.2 ± 24.7
Forests						
500 m	1.5 ± 1.5	8.1 ± 6.2	1.3 ± 1.4	4.9 ± 4.5	3.2 ± 2.1	4.7 ± 3.6
1 km	3.9 ± 3.0	17.8 ± 10.5	3.4 ± 3.0	9.4 ± 7.5	9.2 ± 4.5	10.2 ± 6.2
2.5 km	18.8 ± 14.9	59.7 ± 24.7	15.3 ± 8.1	23.8 ± 12.33	37.8 ± 16.3	**37.6 ± 19.4**
5 km	61.38 ± 42.1	146.6 ± 48.9	49.5 ± 19.6	62.0 ± 21.42	110.3 ± 29.4	**107.7 ± 40.9**
7.5 km	124.4 ± 75.9	223.2 ± 64.3	92.98 ± 31.2	104.7 ± 28.0	184.0 ± 32.4	198.9 ± 63.0
10 km	203.6 ± 114.1	295.9 ± 63.4	139.3 ± 46.0	155.4 ± 38.37	271.1 ± 49.1	306.0 ± 80.8
Heath forests						
500 m	1.2 ± 1.5	8.4 ± 6.1	1.2 ± 1.4	4.8 ± 4.7	2.9 ± 2.2	7.3 ± 5.2
1 km	2.8 ± 2.7	18.4 ± 10.8	3.2 ± 3.5	9.0 ± 7.9	6.7 ± 3.8	17.5 ± 11.4
2.5 km	12.1 ± 12.1	62.3 ± 28.2	12.6 ± 9.2	20.4 ± 13.6	25.9 ± 11.6	61.7 ± 28.0
5 km	41.0 ± 35.2	149.3 ± 55.1	41.0 ± 22.3	49.1 ± 23.8	79.8 ± 35.0	152.6 ± 50.6
7.5 km	83.2 ± 67.8	217.5 ± 67.2	71.9 ± 30.2	76.9 ± 26.3	150.4 ± 58.8	234.0 ± 61.1
10 km	143.3 ± 108.9	275.9 ± 60.5	103.2 ± 36.4	112.1 ± 32.1	226.6 ± 75.5	295.8 ± 62.9
Herb-rich forests						
500 m	0.6 ± 0.9	1.2 ± 1.0	**0.4 ± 0.5**	**0.0 ± 0.0**	1.6 ± 1.6	3.1 ± 3.0
1 km	1.2 ± 0.9	2.0 ± 1.1	**1.4 ± 1.4**	**0.7 ± 0.6**	3.6 ± 2.5	6.3 ± 4.7
2.5 km	2.1 ± 1.2	4.2 ± 2.3	**1.3 ± 1.4**	**0.7 ± 0.6**	12.1 ± 6.6	21.7 ± 9.8
5 km	6.4 ± 1.4	9.6 ± 3.1	**3.0 ± 2.6**	**1.7 ± 1.2**	37.4 ± 14.5	54.0 ± 14.6
7.5 km	11.6 ± 2.6	12.9 ± 1.9	**5.8 ± 3.0**	**5.3 ± 5.1**	71.5 ± 20.0	82.6 ± 14.5
10 km	15.5 ± 4.8	17.3 ± 3.5	**8.1 ± 3.5**	**7.3 ± 4.5**	108.8 ± 26.4	109.0 ± 13.4

(*continued on next page*)

Appendix D (*continued*)

	Area 1	Area 2	Area 3			
	WKHs	Reserves	WKHs	Reserves	WKHs	Reserves
Spruce mires						
500 m	0.7 ± 0.9	3.3 ± 2.3	0.6 ± 0.5	3.3 ± 2.9	0.9 ± 1.1	5.9 ± 4.2
1 km	1.3 ± 0.8	7.6 ± 4.6	0.7 ± 0.5	6.8 ± 5.4	1.6 ± 1.4	15.9 ± 10.8
2.5 km	3.7 ± 1.7	29.0 ± 15.4	5.5 ± 4.3	16.4 ± 12.5	4.8 ± 3.1	54.7 ± 29.5
5 km	14.4 ± 10.7	68.7 ± 26.7	20.4 ± 11.8	23.8 ± 14.9	18.6 ± 8.0	105.1 ± 35.0
7.5 km	29.4 ± 21.1	95.8 ± 27.9	**36.9 ± 11.2**	**30.0 ± 15.2**	43.7 ± 20.9	134.4 ± 31.6
10 km	55.5 ± 41.9	119.1 ± 23.3	**42.8 ± 11.0**	**36.8 ± 15.3**	80.3 ± 43.3	150.8 ± 32.3
Herb-rich spruce mires						
500 m	0.7 ± 0.9	1.2 ± 1.1	–	–	0.8 ± 1.1	2.0 ± 1.7
1 km	0.7 ± 0.9	1.2 ± 1.1	–	–	1.5 ± 1.4	5.0 ± 3.8
2.5 km	2.2 ± 1.1	4.0 ± 1.7	–	–	3.9 ± 2.6	14.1 ± 7.3
5 km	6.0 ± 1.5	9.3 ± 2.2	–	–	12.6 ± 6.7	28.9 ± 6.4
7.5 km	10.2 ± 2.1	10.7 ± 1.7	–	–	24.5 ± 8.5	36.5 ± 4.5
10 km	13.6 ± 3.5	14.5 ± 2.1	–	–	40.6 ± 12.1	46.3 ± 6.0

References

Anonymous, 2008. Finnish Statistical Yearbook of Forestry. Metsäntutkimuslaitos, Helsinki.

Andrén, H., 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos 71, 355–366.

Angelstam, P., Andersson, L., 2001. Estimates of the needs for forest reserves in Sweden. Scand. J. Forest. Res. 3 (Suppl.), 38–51.

Aune, K., Jonsson, B., Moen, J., 2005. Isolation and edge effects among woodland key habitats in Sweden: is forest policy promoting fragmentation? Biol. Conserv. 124, 89–95.

Balmford, A., Whitten, T., 2003. Who should pay for tropical conservation, and how could the costs be met? Oryx 37, 238–250.

Bélisle, M., 2005. Measuring landscape connectivity: the challenge of behavioral landscape ecology. Ecology 86, 1988–1995.

Bender, D.J., Contreras, T.A., Fahrig, L., 1998. Habitat loss and population decline: a meta-analysis of the patch size effect. Ecology 79, 517–533.

Brown, J.H., Kodric-Brown, A., 1977. Turnover rates in insular biogeography: effect of immigration on extinction. Ecology 58, 445–449.

Bunn, A., Urban, D., Keitt, T., 2000. Landscape connectivity: a conservation application of graph theory. J. Environ. Manage. 59, 265–278.

de Nooy, W., Mrvar, A., Batagelj, V., 2005. Exploratory Social Network Analysis with Pajek. Cambridge University Press.

Edman, M., Kruys, N., Jonsson, B.G., 2004a. Local dispersal sources strongly affect colonization patterns of wood-decaying fungi on spruce logs. Ecol. Appl. 14, 893–901.

Edman, M., Gustafsson, M., Stenlid, J., Jonsson, B.G., Ericson, L., 2004b. Spore deposition of wood-decaying fungi: importance of landscape composition. Ecography 27, 103–111.

Esseen, P., Ehnström, B., Ericsson, L., Sjöberg, K., 1997. Boreal forests. Ecol. Bull., 16–47.

Estrada, E., Bodin, O., 2008. Using network centrality measures to manage landscape connectivity. Ecol. Appl. 18, 1810–1825.

Fahrig, L., Merriam, G., 1985. Habitat patch connectivity and population survival. Ecology 66, 1762–1768.

Fahrig, L., Paloheimo, J., 1988. Effect of spatial arrangement of habitat patches on local-population size. Ecology 69, 468–475.

Franken, R.J., Hik, D.S., 2004. Influence of habitat quality, patch size and connectivity on colonization and extinction dynamics of collared pikas *Ochotona collaris*. J. Anim. Ecol. 73, 889–896.

Freeman, L., Borgatti, S., White, D., 1991. Centrality in valued graphs: a measure of betweenness based on network flow. Soc. Networks 13, 141–154.

Gustafsson, L., Hylander, K., Jacobson, C., 2004. Uncommon bryophytes in Swedish forests—key habitats and production forests compared. Forest Ecol. Manag. 194, 11–22.

Hanski, I., 2006. Täsmäsuojelun mahdollisuus – vai mahdottomuus. In: Jalonen, R. et al. (Eds.), Uusi Metsäkirja. Gaudeamus, Helsinki, pp. 200–207.

Hanski, I., Kuussaari, M., Nieminen, M., 1994. Metapopulation structure and migration in the butterfly *Melitaea Cinxia*. Ecology 75, 747–762.

Hansson, L., 2001. Key habitats in Swedish managed forests. Scand. J. Forest Res. 3 (Suppl.), 52–61.

Johansson, P., Ehrlén, J., 2003. Influence of habitat quantity, quality and isolation on the distribution and abundance of two epiphytic lichens. J. Ecol. 91, 213–221.

Johansson, P., Gustafsson, L., 2001. Red-listed and indicator lichens in woodland key habitats and production forests in Sweden. Can. J. Forest Res. 31, 1617–1628.

Jonsson, M., 2003. Colonisation ability of the threatened tenebrionid beetle *Oplocephala haemorrhoidalis* and its common relative *Bolitophagus reticulatus*. Ecol. Entomol. 28, 159–167.

Keitt, T.H., Urban, D.L., Milne, B.T., 1997. Detecting critical scales in fragmented landscapes. Conservation Ecology 1 (1), 4.

Knight, R.L., 1999. Private lands: the neglected geography. Conserv. Biol. 13, 223–224.

Korhonen, K.T., Ihalainen, A., Heikkinen, J., Henttonen, H., Pitkänen, J., 2007. Suomen metsävarat metsäkeskuksittain 2004–2006 ja metsävarojen kehitys 1996–2006. Metsätieteen Aikakauskirja 2B, 149–213.

Meriluoto, M., Soininen, T., 1998. Metsäluonnon Arvokkaat Elinympäristöt. Metsälehti Kustannus, Helsinki.

Moilanen, A., Nieminen, M., 2002. Simple connectivity measures in spatial ecology. Ecology 83, 1131–1145.

Mönkkönen, M., Reunanen, P., 1999. On critical thresholds in landscape connectivity: a management perspective. Oikos 84, 302–305.

Nilsson, C., Gotmark, F., 1992. Protected areas in Sweden: is natural variety adequately represented? Conserv. Biol. 6, 232–242.

Östlund, L., Zackrisson, O., Axelsson, A.L., 1997. The history and transformation of a Scandinavian boreal forest landscape since the 19th century. Can. J. Forest Res. 27, 1198–1206.

Pascual-Hortal, L., Saura, S., 2008. Integrating landscape connectivity in broad-scale forest planning through a new graph-based habitat availability methodology: application to capercaillie (*Tetrao urogallus*) in Catalonia (NE Spain). Eur. J. Forest Res. 127, 23–31.

Pascual-Hortal, L., Saura, S., 2006. Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landscape Ecol. 21, 959–967.

Penttila, R., Lindgren, M., Miettinen, O., Rita, H., Hanski, I., 2006. Consequences of forest fragmentation for polyporous fungi at two spatial scales. Oikos 114, 225–240.

Perhans, K., Gustafsson, L., Jonsson, F., Nordin, U., Weibull, H., 2007. Bryophytes and lichens in different types of forest set-asides in boreal Sweden. Forest Ecol. Manage. 242, 374–390.

Pressey, R.L., 1994. Ad hoc reservations: forward or backward steps in developing representative reserve systems? Conserv. Biol. 8, 662–668.

Pykälä, J., 2004. Effects of new forestry practices on rare epiphytic macrolichens. Conserv. Biol. 18, 831–838.

Pykälä, J., 2007. Implementation of Forest Act habitats in Finland: does it protect the right habitats for threatened species? Forest Ecol. Manage. 242, 281–287.

Ranius, T., Hedin, J., 2001. The dispersal rate of a beetle, *Osmoderma eremita*, living in tree hollows. Oecologia 126, 363–370.

Rassi, P., Alanen, A., Kanerva, T.,Mannerkoski, I. (Eds.), 2001. Suomen lajien uhanalaisuus 2000. Uhanalaisten lajien II seurantaryhmä. Ympäristöministeriö, Suomen ympäristökeskus, Edita.

Saunders, D.A., Hobbs, R.J., Margules, C.R., 1991. Biological consequences of ecosystem fragmentation: a review. Conserv. Biol. 5, 18–32.

Saura, S., Torné, J., 2009. Conefor Sensinode 2.2: a software package for quantifying the importance of habitat patches for landscape connectivity. Environ. Modell. Softw. 24, 135–139.

Selonen, V.A.O., Kotiaho, J.S. 2006. Suojavyöhykkeen leveyden ja muodostumisajan vaikutus erityisen tärkeiden puroelinympäristöjen monimuotoisuuteen. In: Horne, P., Koskela, T., Kuusinen, M., Otsamo, A., Syrjänen, K. (Eds.), Metson jäljillä. Etelä-Suomen metsien monimuotoisuusohjelman tutkimusraportti. Maa- ja metsätalousministeriö, Ympäristöministeriö, Metsäntutkimuslaitos and Suomen ympäristökeskus, pp. 327–329.

Sillett, S.C., McCune, B., Peck, J.E., Rambo, T.R., Ruchty, A., 2000. Dispersal limitations of epiphytic lichens result in species dependent on old-growth forests. Ecol. Appl. 10, 789–799.

Sippola, A., Mönkkönen, M., Renvall, P., 2005. Polypore diversity in the herb-rich woodland key habitats of Koli National Park in eastern Finland. Biol. Conserv. 126, 260–269.

Stokland, J.N., 1997. Representativeness and efficiency of bird and insect conservation in Norwegian boreal forest reserves. Conserv. Biol. 11, 101–111.

Taylor, P.D., Fahrig, L., Henein, K., Merriam, G., 1993. Connectivity is a vital element of landscape structure. Oikos 68, 571–573.

Tischendorf, L., Fahrig, L., 2000. On the usage and measurement of landscape connectivity. Oikos 90, 7–19.

Urban, D., Keitt, T., 2001. Landscape connectivity: a graph-theoretic perspective. Ecology 82, 1205–1218.

Urban, D.L., Minor, E.S., Treml, E.A., Schick, R.S., 2009. Graph models of habitat mosaics. Ecol. Lett. 12, 260–273.

Uusitalo, A., Paakkolanvaara, J., 2007. Kylien kaunokit, soiden sarat. Keski-Suomen uhanalaiset kasvit. Keski-Suomen ympäristökeskus.

van Langevelde, F., 2000. Scale of habitat connectivity and colonization in fragmented nuthatch populations. Ecography 23, 614–622.

Verboom, J., Schotman, A., Opdam, P., Metz, J.A.J., 1991. European nuthatch metapopulations in a fragmented agricultural landscape. Oikos 61, 149–156.

Virkkala, R., 1996. Metsien suojelualueverkon rakenne ja kehittämistarpeet: ekologinen lähestymistapa. Suomen Ympäristö 16, Suomen ympäristökeskus, Helsinki.

Walters, S., 2007. Modeling scale-dependent landscape pattern, dispersal, and connectivity from the perspective of the organism. Landscape Ecol. 22, 867–881.

With, K.A., Gardner, R.H., Turner, M.G., 1997. Landscape connectivity and population distributions in heterogeneous environments. Oikos 78, 151–169.

Yrjönen, K., 2004. Metsälain erityisen tärkeät elinympäristöt. Kartoitus yksityismetsissä 1998–2004. Loppuraportti. Maa-ja metsätalousministeriö. MMM:n julkaisuja 9 (2004), 1–60.

II

GRAPH-THEORETIC CONNECTIVITY MEASURES: WHAT DO THEY TELL US ABOUT CONNECTIVITY?

by

Anne Laita, Janne S. Kotiaho & Mikko Mönkkönen 2011

Landscape Ecology 26: 951-967.

Landscape Ecol (2011) 26:951–967
DOI 10.1007/s10980-011-9620-4

RESEARCH ARTICLE

Graph-theoretic connectivity measures: what do they tell us about connectivity?

A. Laita · J. S. Kotiaho · M. Mönkkönen

Received: 9 March 2010 / Accepted: 28 May 2011 / Published online: 12 June 2011

Abstract Graph-theoretic connectivity analyses have received much attention in connectivity evaluation during the last few years. Here, we explore the underlying conceptual differences of various graph-theoretic connectivity measures. Based on connectivity analyses from three reserve networks in forested landscapes in Central Finland, we illustrate how these conceptual differences cause inconsistent connectivity evaluations at both the landscape and patch level. Our results also illustrate how the characteristics of the networks (patch density) may affect the performance of the different measures. Many of the connectivity measures react to changes in habitat connectivity in an ecologically undesirable manner. Patch prioritisations based on a node removal analysis were sensitive to the connectivity measure they were based on. The patch prioritisations derived from different measures showed a disparity in terms of how much weight they put on patch size versus patch location and how they value patch location. Although graphs operate at the interface of structure and function, there is still much to do for incorporating the inferred ecological process into graph structures and analyses. If graph analyses are going to be used for real-world management and conservation purposes, a more thorough understanding of the caveats and justifications of the graph-theoretic connectivity measures will be needed.

Keywords Functional connectivity · Graph theory · Reserve network · Component · Patch prioritisation

Introduction

Habitat loss and fragmentation pose two primary threats to biodiversity across spatial scales that range from the global to very local ones. Fragmentation confounds and intensifies the effect of pure habitat loss when the amount of habitat falls below a critical threshold (Andren 1994; Mönkkönen and Reunanen 1999). Although habitat loss and fragmentation are separate components affecting the patterning of habitat (i.e. habitat amount and its configuration in a landscape), their independent roles are difficult to evaluate (Fahrig 1997,2003; Bender et al. 1998; Wiegand et al. 2005).

The intertwined ecological consequences of habitat loss and fragmentation can be understood and measured on the basis of the concept of connectivity. Connectivity supports ecological flows in a landscape and, with various mechanisms, influences the viability of spatially structured populations. Habitat connectivity is needed to sustain spatially dependent ecological processes, and it is a necessity for the long-term persistence of biodiversity (Fahrig and

A. Laita (✉) · J. S. Kotiaho · M. Mönkkönen
Department of Biological and Environmental Science,
University of Jyväskylä, PO Box 35, Survontie 9,
Jyväskylä 40014, Finland
e-mail: anne.laita@jyu.fi

Merriam 1994; Hanski 1999). Connectivity conservation can also be considered to be a prerequisite for ecologically and economically efficient conservation practices.

The most commonly used definition for landscape connectivity is rooted in a functional connectivity concept, and it is described as 'the degree to which landscape facilitates or impedes movement of organisms among resource patches' (Taylor et al. 1993). Underlying the functional connectivity concept is the idea that the connectivity experienced by an organism is result of the behavioural responses of the organisms to physical landscape structure (e.g. Tischendorf and Fahrig 2000a, b; Bélisle 2005; Kindlmann and Burel 2008). However, all of the definitions of functional connectivity are conceptually so broad that they leave room for a range of interpretations and applications. The very nature of the concept of connectivity has been approached from several (also conflicting) perspectives during the past decade (Tischendorf and Fahrig 2000a, 2001; Moilanen and Hanski 2001; Goodwin 2003), which indicates that the definition of the concept is far from simple, and that the focuses of the researcher are reflected in it.

Graphs are versatile models for analysing a wide range of practical problems concerned with the properties and functions of networks (Gross and Yellen 2006). In landscape ecology, graphs are abstractions of landscapes where habitat patches are represented as spatially explicit nodes and functional connections between the nodes as links (Fall et al. 2007; for in-depth graph definitions in the field of landscape ecology we recommend Bunn et al. 2000, Urban and Keitt 2001; Urban et al. 2009). Graph theory, with its algorithms, has given rise to many connectivity measures with varying degrees of complexity and differing underlying assumptions. Some connectivity measures have been adopted to landscape ecology from the general graph definitions and methodology of other disciplines, while other measures were specifically designed for the evaluation of landscape connectivity. Graph metrics that were developed for the uses of other disciplines may not, however, be suitable for the evaluation of landscape connectivity due to its very special characteristics and needs (Saura and Rubio 2010).

Graphs operate at the interface of structure and process (e.g. Urban and Keitt 2001; Urban et al. 2009). In landscape ecology, graph structures are defined in reference to the dispersal ability of a species and can be analysed as structures as such or with a specific relevance to the underlying process, such as gene flow, flux of dispersing individuals (Urban and Keitt 2001; Minor and Urban 2007a, b), species occurrence (Andersson and Bodin 2009) or species invasions (Ferrari and Lookingbill 2009). Graph-theoretic connectivity measures vary in terms of how they infuse and deal with the underlying ecological process, although they are ultimately founded on the concept of metapopulation with spatially interrelated subpopulations (Urban et al. 2009).

Graph-theoretic connectivity analyses lend themselves to functional connectivity evaluations both at the level of the entire networks and at the level of individual patches. At the network level, they have been used to evaluate network connectedness for a focal species (Keitt et al. 1997; Bunn et al. 2000), in the design of reserve network (Fuller et al. 2006) and in connectivity conservation for the habitat of threatened species (Fall et al. 2007). Connectivity evaluation at the level of individual patches is often called patch prioritisation, because in many cases it is used to select the most valuable patches in a habitat network (Jordán et al. 2003; Rothley and Rae 2005; Pascual-Hortal and Saura 2008). Patch prioritisation has also been used to extract configurational properties of patches into a variable that can be used in statistical analyses to explain ecological process (e.g. species occurrence, colonisation probability) (van Langevelde 2000) or to reveal the roles of patches in a landscape (Keitt et al. 1997; Bunn et al. 2000; Urban and Keitt 2001).

The graph-theoretic connectivity measures do not form a single story about connectivity, but represent a full spectrum of specific measures that capture different aspects of connectivity. This is important to remember when the measures are selected for different kinds of applications and their outcomes interpreted. Pascual-Hortal and Saura (2006) and Saura and Pascual-Hortal (2007) have made a valuable effort in the systematic comparison of different measures and their properties. They have investigated the performance of measures in landscapes varying in their level of habitat loss and fragmentation with an emphasis on the outcome (whether it is desirable or not) rather than on the mechanisms that produced the outcome. Although empirical data sets on the movements of individuals (i.e. realised connectivity) represent the "truth" against which the performance of connectivity measures should be evaluated and calibrated, connectivity measures can also, to some extent, be evaluated and understood on theoretical grounds. Theoretical evaluation is especially suitable for revealing possible unwanted behaviour of the measures (that is not in agreement with the concept of connectivity), and comparisons between measures may provide insights into their characteristics and performance.

Here, we shall first review and compare the available graph-theoretic connectivity measures in terms of their conceptual differences. One of the major factors leading to differences among the measures is whether connectivity can only "prevail" between habitat patches or if intrapatch connectivity is also acknowledged. For instance, what is the connectivity for a landscape that is fully covered by habitat? Zero or the maximum? The non-acceptance that area within a habitat patch may contribute to connectivity may lead to the counterintuitive outcome that connectivity has a positive relationship with fragmentation (Tischendorf and Fahrig 2000b; Pascual-Hortal and Saura 2006; Saura and Pascual-Hortal 2007; Matisziw and Murray 2008). The relationship of the measure to intrapatch connectivity also largely determines whether the measure explicitly tracks the amount of suitable habitat in a landscape. Although the way how the pairwise distances between patches are determined (e.g. as Euclidean, cost-modified distances or based on simulation) is independent of the connectivity measure, how the connections between patches are defined (e.g. among all pairs of patches or not, as direct connections only or acknowledging paths) can differ widely among measures with anticipated influence on their behaviour.

In our review, we shall briefly describe a number of graph-theoretical measures that are used in the evaluation of landscape connectivity. We shall use simple examples to illustrate what kind of counterintuitive connectivity evaluations the measure may possibly produce and with which mechanisms. We draw on a connectivity concept that recognises both habitat amount and connections among and inside patches as contributing to connectivity (the so-called habitat availability concept described by Saura and Pascual-Hortal 2007; and Pascual-Hortal and Saura; 2008). In contrast to the original definition of functional connectivity by Taylor et al. (1993), this concept acknowledges within resource patch connectivity and recognises habitat amount as a critical factor contributing to dispersal among resource patches. The sizes of a source and a target patch are linked with the dispersal probability between the patch pair within a given time period (more dispersing individuals with increasing donating area; growing chance of ending in a target patch with increasing number of dispersing individuals, and with the increasing size of a target patch). The connectivity based on this concept may be briefly described as the amount of habitat that is available to a species dispersal (given its assumed dispersal ability) at the landscape level. The concept can easily be extended to dispersal flux if the habitat area is assumed to scale linearly with the number of migrants. The connectivity rises if (i) for a given amount of habitat, the connection status is improved or (ii) for a given connection status, the amount of habitat increases. The connection status is improved when the degree of linkage among patches increases and more so when either the strength of connections among the patches increases (so that proximal patches contribute more to connectivity than distant ones) or the habitat area that is linked increases.

In the second part of the article, we analyse empirical data from three reserve networks located in forested landscapes in Central Finland to illustrate how the underlying conceptual differences in the connectivity measures influence the connectivity ratings at the network and patch level. At the network level, we compare the performance of the connectivity measures in terms of how they react when the existing forest reserve network is augmented with small-sized woodland key habitats (WKHs). The numerical value of a connectivity measure should, if lined with the connectivity concept, react positively to the addition of WKHs into the reserve network, as both habitat area and the number (and strength) of interpatch connections increase. At the patch level, we compare the prioritisations obtained from different measures.

Review of the graph-theoretic connectivity measures

We classify the graph-theoretic connectivity measures broadly into two groups: network coherence measures and flux measures (which corresponds to the division of the measures into indirect and direct ones by Matisziw and Murray 2008). The coherence measures provide information about some element of network structure (in respect to coherence or fragmentation in reverse) taking into account species dispersal abilities. Coherence measures usually react strongly to the component structure of the network, acknowledging a 'single component' structure as the one maximising connectivity with a given amount of habitat. In contrast, the flux measures summarise interpatch connections between all pairs of patches. The flux measures can either take into account direct interpatch connections only (direct links) or also allow paths, i.e. indirect links via stepping stone patches. The flux measures mainly aim to evaluate the rate of flow of dispersing individuals at the level of the entire habitat network.

We have summarised in Table 1 the various connectivity measures that have been used in landscape ecological connectivity analyses, and below we shall discuss in detail the ones that are most commonly used.

Network coherence measures

Graph diameter is the longest path between any two habitat patches in the graph, where the path length between the patches is itself the shortest possible path (Bunn et al. 2000). As Ferrari et al. (2007) stated, 'a large graph diameter can either be positive or negative [*for an individual traversing a landscape*] and needs to be interpreted with caution'. Graph diameter is purely a topological measure, which does not take patch areas into account, and generally increases with the increasing number of patches brought about by fragmentation. Graph diameter is usually calculated for the largest component, and within that component, it only provides information on the connectivity between a single pair of patches. *Characteristic path length*, CPL (or average path length) is the average of the shortest path lengths calculated between all pairs of reachable patches in a network. According to Minor and Urban (2007a, b), 'if CPL is short, all patches tend to be easily reachable'. As the measure concentrates only on reachable pairs of patches, CPL is also short when a network is composed of isolated, internally well-connected components

Table 1 Summary of the connectivity measures used in graph-theoretic connectivity analyses

Measure			References
Network coherence measures			
Characteristic path length, CPL		B	Minor and Urban (2007)
Coincidence probability (class/landscape, CCP/LCP)	C	B	Pascual-Hortal and Saura (2006)
Correlation length	C	B	Keitt et al. (1997), Rothley and Rae (2005)
Expected cluster size, ECS	C	B	O'Brien et al. (2006), Fall et al. (2007)
Graph diameter	C	B	Bunn et al. (2000), Bodin and Norberg (2007), Ferrari et al. (2007), Minor and Urban (2007)
Graph-derived proportional measures, e.g.,			
Ratio of graph diameter to the size of the largest component	C	B	Minor and Urban (2007)
Ratio of the proportion of habitat in the largest patch to the proportion of habitat in the largest cluster (F-measure)	C	B	Ferrari et al. (2007), Lookingbill et al. (2010)
Flux measures			
(Area-weighted) flux (AW)F		P	Bunn et al. (2000), Urban and Keitt (2001)
Harary index, H		B	Jordán et al. (2003)
Habitat availability indices			
Integral index of connectivity, IIC		B	Pascual-Hortal and Saura (2006), Pascual-Hortal and Saura (2008)
Probability of connectivity, PC		P	Saura and Pascual-Hortal (2007)
Total network connectivity		P	Matisziw and Murray (2008)

Measures based on component approach are marked with C. Binary and probabilistic connectivity models are marked with B and P, respectively (comprised possibly of few patches). Thus, fragmentation, in the form of isolated components, may lead to a false indication of connectivity. In the case of a single component network, CPL reaches a minimum value of one when all patches are within one step from each other—irrespective of the number of patches and habitat amount involved.

Correlation length measures the average distance an individual with a given dispersal ability can move before reaching a barrier (Keitt et al. 1997). Correlation length is calculated as an area-weighted mean radius of gyration of all the components in a landscape as follows (for raster data):

$$C_d = \frac{\sum_{i=1}^{\mathrm{NC}} n_i . R_i}{\sum_{i=1}^{\mathrm{NC}} n_i},$$

where NC is the number of components in the landscape, n_i is the number of habitat cells (i.e. pixels) in component i, and:

$$R_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \sqrt{\left(x_j - \overline{x_i}\right)^2 + \left(y_j - \overline{y_i}\right)^2},$$

where R_i is the radius of gyration of component i, $\overline{x_i}$ and $\overline{y_i}$ are the mean x and y coordinates of the habitat cells in component i, and x_{j} and y_j are the coordinates of the jth habitat cell in component i.

Similar to the graph diameter, there is the undesired property in the correlation length that it is positively related to increasing fragmentation. For example, two connected patches score higher in correlation length than one patch with the same total area. Increasing interpatch distances (within the dispersal ability of a species) also increase the value of this measure, although any species would likely benefit from the proximity of the patches. Correlation length does not explicitly account for the total habitat area, because the habitat area is masked by the extent and shape of a component. Habitat area in a component can be small, but if distributed evenly within a component, it can still provide an opportunity for high average dispersal distances.

Expected cluster size (first introduced by O'Brien et al. 2006) represents an area-weighted mean cluster (i.e. component) size calculated as:

$$\mathrm{ECS} = \frac{\sum_{i=1}^{\mathrm{NC}} a_i^2}{a},$$

where NC is the number of components in the landscape, a_i is the habitat area in component i and a is the total habitat area over all components of the graph.

ECS is the size of the component in which a point randomly located within a habitat area is expected to reside at a given threshold distance w. ECS carries information on the amount of habitat within a component, but it still does not react ecologically meaningfully to the amount of habitat in a landscape. For example, the value of ECS increases with the loss of isolated patches/components with a small habitat area, although the total habitat area in the landscape diminishes.

Landscape coincidence probability, LCP, is the probability that two points located randomly within a landscape reside in the same habitat component (Pascual-Hortal and Saura 2006). It is computed as:

$$\mathrm{LCP} = \sum_{i=1}^{\mathrm{NC}} \left(\frac{c_i}{A_{\mathrm{L}}}\right)^2,$$

where NC is the number of components in the landscape, c_i is the sum of the sizes of all the patches belonging to component i and A_{L} is the total landscape area.

LCP is reactive to the amount of habitat in the landscape and shows a decrease with increasing fragmentation. It evaluates (like ECS) reachability between patches as defined by the component structure only, but does not provide information about the internal connectivity of components (discussed in more detail in 'Discussion' section).

Flux measures

Flux measures can be based on a binary or a probabilistic connection model. The probabilistic connection model weights the links with the dispersal probability between two habitat patches. The probability of direct dispersal between patches is determined on the basis of a chosen dispersal kernel, most often a decreasing exponential function of the interpatch distance:

$$p_{ij} = e^{-k \cdot d_{ij}},$$

where k is a constant and d_{ij} is the distance (Euclidean or effective distance) between patches i and j.

Area-weighted flux, AWF, evaluates the area-weighted flux between all pairs of patches:

$$\mathrm{AWF} = \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} p_{ij} \cdot a_i \cdot a_j,$$

where n is the total number of patches, p_{ij} is the probability of direct dispersal between patches i and j, and a_i and a_j are the sizes of patches i and j.

AWF does not account for the dispersal potential within a patch itself as it concentrates only on the flux between patches. For this reason, it does not react ecologically meaningfully to the amount of habitat in a landscape; for example, it neglects the loss of an isolated patch ($p_{ij} = 0$) regardless of its size. It also reacts in an undesirable way to fragmentation as, for example, the connectivity value increases with an increasing number of habitat patches when the habitat area and interpatch distances are controlled for. The connection model of the measure does not allow for indirect interpatch connections mediated by stepping stone patches.

Connectivity measures rooted in habitat availability (Integral index of connectivity and Probability of connectivity, presented below) integrate the habitat area existing within patches with the area made available by the interpatch connections into a single measure (Pascual-Hortal and Saura 2006, 2008; Saura and Pascual-Hortal 2007). If habitat patch area (or other patch attribute) is used as a surrogate for the number of dispersing individuals, habitat availability measures are easily interpreted as flux measures. Habitat availability measures avoid undesired responses to increased fragmentation and habitat loss.

Integral index of connectivity, IIC, is a habitat availability measure with a binary connection model (Pascual-Hortal and Saura 2008):

$$\mathrm{IIC} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_i \cdot a_j}{1 + l_{ij}}}{A_{\mathrm{L}}^2},$$

where n is the total number of patches, a_i and a_j are the sizes of patches i and j, l_{ij} is the number of links in the shortest path between patches i and j (and equals zero for $i = j$), and A_{L} is the total landscape area. IIC reaches unity when the landscape is occupied by the given habitat.

As shown by Pascual-Hortal and Saura (2006), IIC can reliably provide information on habitat amount and the degree of connectivity between patches, but with the restrictions brought about by its binary view of connectivity (Saura and Pascual-Hortal 2007). IIC evaluates the strength of connections between patches in a topological sense (i.e., the inverse of 1 + the number of links in the shortest path separating the patches), and thus recognises the increasing topological distances between patches as lower connectivity. IIC favours habitat located in a single large patch (in which the habitat area is separated by 0 links), whereas patches with direct or indirect connections (a path length of one link or more) are downscaled by the increasing denominator. For example, two directly connected patches would have lower connectivity than one large patch of the same total area.

Probability of connectivity, PC, measure (Saura and Pascual-Hortal 2007) is calculated as:

$$\mathrm{PC} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_i \cdot a_j \cdot p_{ij}^{*}}{A_{\mathrm{L}}^2},$$

where n is the total number of patches, a_i and a_j are the sizes of patches i and j, and A_{L} is the total landscape area. p_{ij}^{*} is defined as the maximum product probability of all possible paths between patches i and j. Product probability of a path is the product of all the link weights (p_{ij}) included in the path. For patches close enough, p_{ij}^{*} is reduced to the direct dispersal probability p_{ij}, but for more distant patches the 'best path' passes through stepping stone patches. When $i = j$, the p_{ij}^{*} equals 1.

It has been recommended that PC be used as a connectivity measure, because it is, according to Saura and Pascual-Hortal (2007), the only measure having all the properties of an ideal connectivity measure. It reacts meaningfully to habitat loss and network fragmentation. PC also possesses the richest connection model of the measures in widespread use in connectivity evaluation. PC does not, however, account for other connections between any two patches besides the most probable path (i.e. it does not react to the number of connections between the patches).

Empirical comparison of the performance of the graph-theoretic connectivity measures: habitat network analyses in forested landscape in Finland

With our empirical analyses, we wanted to address the following questions: (i) How do different graph-theoretic connectivity measures perform when the existing reserve network is augmented with small-sized WKHs? (Network level), (ii) How consistently do patch prioritisations based on different measures value patches, and how sensitive are these prioritisations to the density of habitat patches in a landscape? (Patch level). Patch density may have important consequences for the ability of measures to consistently value patches because network properties and the roles of patches in sparse networks differ in an anticipated way from those of denser networks. We used data from three habitat networks, which were composed of reserve patches of a given habitat type added with WKH patches of the same habitat type. The habitat types were herb-rich forests, spruce mires and heath forests. The total landscape area (matrix and habitat area included) in all the networks was ca. 500 km^2. Networks varied more than 10-fold in terms of the patch number (i.e. graph order) and density (number of patches per landscape area); the herb-rich forest network had 42 (17 reserves; 25 WKHs), the spruce mire network 158 (134 reserves; 24 WKHs) and the heath-forest network 528 (242 reserves; 286 WKHs) patches. The total habitat area in the herb-rich forest network was 79 ha (36% of the area in WKHs), in the spruce mire network 283 ha (10% of the area in WKHs) and in the heath-forest network 2634 ha (6% of the area in WKHs) representing more than 30-fold difference in habitat area. The networks were located in Central Finland (61°26′–63°37′N, 24°6′–26°46′E); the herb-rich forest and spruce mire networks were in the northern parts of the area, and the heath-forest network was in the south. The networks are described in more detail in Laita et al. (2010).

Overall network connectivity measures were calculated with a threshold distance ranging from 200 m to 25 km. Patch importance measures were calculated at the thresholds distances of 500 m, 1 km, 2.5 km, 5 km, 7.5 km and 10 km. Link weights were determined based on Euclidean edge-to-edge distances between patches. For overall network connectivity analyses, we formed two separate networks for each habitat type: one composed of reserve patches only and another network with WKHs added. We plotted the connectivity of both networks to the same graph to see how the measure reacts to the addition of WKHs over the range of threshold distances. The network measures used for overall network analyses were: correlation length; expected cluster size (ECS); landscape coincidence probability (LCP); area-weighted flux (AWF); integral index of connectivity (IIC) and probability of connectivity (PC). For probabilistic measures, AWF and PC, we defined a dispersal probability of 0.05 to correspond to the threshold distance (i.e. the cut-off dispersal distance, if exceeded a patch pair is not connected by a link) used in analyses based on the binary connection model.

For patch prioritisations, we determined how consistently the different connectivity measures value individual patches. Patch prioritisations were calculated with a node removal analysis; each patch was systematically removed from the network, and the reconstructed network was compared with the original network which included the patch (Keitt et al. 1997). The connectivity loss caused by the removal of a patch measures the contribution of the patch to the network connectivity. We compared patch prioritisations for four measures that account for patch area, two of them binary (IIC and LCP) and two probabilistic (AWF and PC). We analysed the proportion of patch importance that was explained by patch area. The proportion was given by the coefficient of determination (R^2) of the regression of the log-transformed patch prioritisation values against the log-transformed patch area. To see how consistently prioritisations derived from different measures value patches, we calculated Pearson's correlations and partial correlations (controlling for patch area) for each measure pair as a function of threshold distance. Correlations were calculated for log-transformed variables. Partial correlations show how consistently different measures value the locations of patches (with the assumption that the part of the prioritisation not explained by patch area represents its locational value in a network). We also calculated full and partial correlations for prioritisations including only the 20 top-ranking patches based on the IIC values averaged over all threshold distances. IIC does not naturally represent any standard, but being rather widely applied it offers an interesting reference for the performance of other measures. By concentrating on the "top-20" patches, we determined whether measures value more consistently the top patches rather than all patches in their prioritisations as the prioritisations are usually used for recognising the most valuable patches.

Graphs were analysed with Conefor Sensinode v.2.2 (Saura and Torné 2009) (for AWF, IIC, LCP and PC) and Spatially Explicit Landscape Event Simulator v.3.3 SELES (Fall and Fall 2001) (for correlation length and ECS).

Results

As expected on the basis of the review above, connectivity evaluations of the connectivity measures differed from each other (Fig. 1). Correlation length, expected cluster size and LCP levelled-off when patches occured as one component. This happened, for example, in the spruce mire network with WKHs when species' dispersal ability exceeded 3 km. After the levelling-off, these network coherence measures did not show further changes with improving dispersal ability. IIC, AWF and PC, however, showed a rise in connectivity with increasing threshold distance even after the network is composed of one component. It remains that IIC levelled-off after all patches are in direct connection with each other. This happened when the threshold distance corresponded to the maximum distance between the pairs of patches. Probabilistic measures (AWF and PC) react in a similarly, despite the underlying differences in their connection models. Both measures showed a rise in connectivity with increasing threshold distance until the probability of dispersal between all pairs of patches reached 1.

Expected cluster size and correlation length did not react in line with the connectivity concept when adding WKHs to the network (Fig. 1). At small threshold distances, they indicated lower connectivity for networks with WKHs than for reserves-only networks. At small threshold distances, the WKH patches increased the number of components in the landscape and, being small-sized, decreased the average component size indicated by expected cluster size (habitat area of component) and correlation length (component extent). LCP, IIC, AFW and PC all recognised that WKHs contributed to increased connectivity, yet differently. For example, probabilistic measures showed a considerably smaller contribution of WKHs to connectivity than binary measures. How WKHs affected connectivity with increasing threshold distance also differed among measures. Binary measures showed a rather rapid rise in connectivity at the threshold distances of ca. 200 m to 5 km. For IIC, this rise with threshold distance was smoother than for other binary measures, which showed a stepwise pattern. Probabilistic measures did not show any range of threshold distances as being critical for connectivity, but expressed a steady rise in connectivity with threshold distance.

Connectivity measures showed different patch prioritisation performances in terms of how much emphasis they put on the effect of patch configuration versus patch size at different threshold distances (Fig. 2). For AWF, the importance of patch size to patch prioritisations rose steadily with increasing threshold distance. The PC measure emphasised, more than AWF, the value of patch size at small threshold distances in patch prioritisations. With IIC and LCP, the relative effect of patch size increased pronouncedly with the threshold distance until the patch importance values were totally determined by patch size. In the herb-rich forest and the spruce mire networks, however, the effect of patch size on patch prioritisations for the two binary measures dropped at intermediate distances (at ca. 1–2.5 km), suggesting the importance of patch configuration at this scale.

Network density had an influence on the relative importance of patch size over patch configuration. In the lowest density network (herb-rich forest), patch importance was determined to a greater extent by patch size at small threshold distances, compared to the other two networks of larger density. In the herb-rich forest network, patch size explained over 60% of patch importance values across all threshold distances for prioritisations based on IIC, LCP and PC. In the other two networks, patch size accounted less for patch priority, ca. 30–60% at small threshold distances. In the herb-rich forest network, the three measures weighted consistently the relative role of patch size on prioritisations across threshold distances. In the spruce mire and the heath-forest networks, the binary, IIC and LCP, and probabilistic PC measures diverged from each other in performance; binary measures put clearly more emphasis on the role of patch size at threshold distances greater than 2.5 km.

The patch prioritisations based on different measures correlated highly in the networks (Fig. 3). This is because all of the prioritisations were to a high degree driven by patch sizes (Fig. 2). The correlations thus mainly reflected the characteristics of

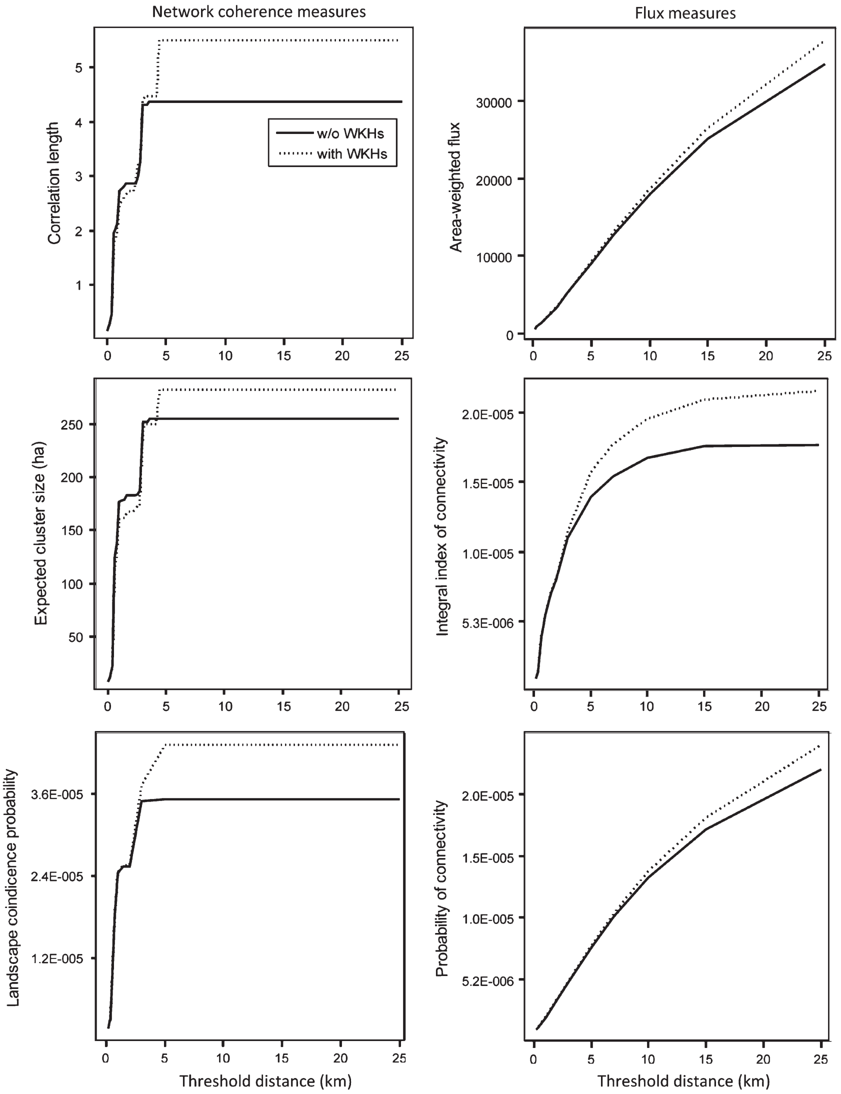


Fig. 1 Comparisons of measure performances as a function of threshold distance for the reserve network of spruce mires with and without woodland key habitats (WKHs)

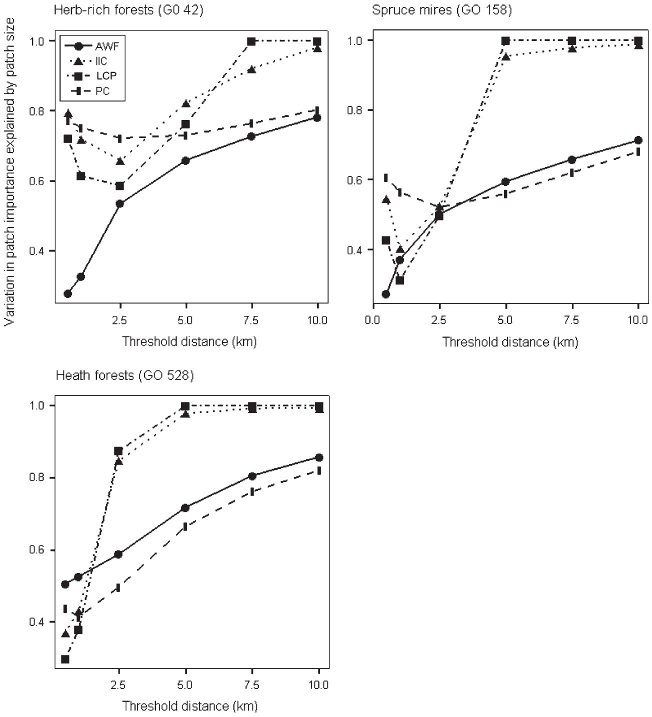


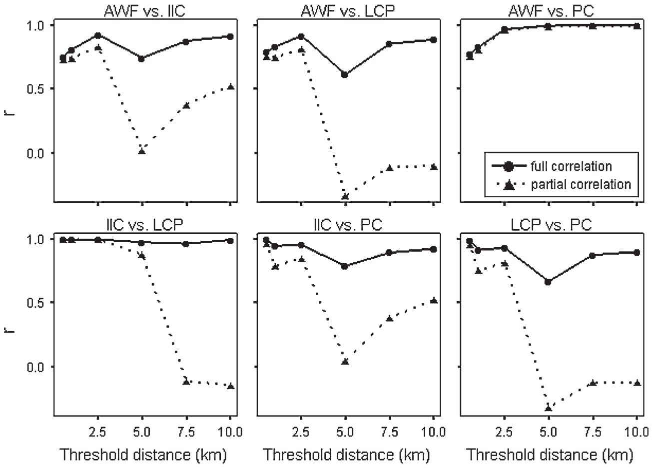
Fig. 2 Comparisons of four connectivity measures (AWF, IIC, LCP, and PC) for the effect of patch size on the patch prioritisation values at different threshold distances for three habitat networks that vary in terms of their graph order (GO). The effect of patch size is calculated by regressing log-transformed patch importance values against log-transformed patch areas and reported as a coefficient of determination (R^2)

measures in terms of how much they value patch size in their prioritisations. When the effect of patch size on the correlation was controlled for, the measures showed differences in their evaluations of the influence of patch location. The two probabilistic measures, AWF and PC, evaluate consistently the configuration component of prioritisations. LCP-valued locations of patches in a different manner compared to other measures at great threshold distances (>5 km). There even existed negative correlations for the prioritisations (AWF vs. LCP, LCP vs. PC and IIC vs. LCP) indicating that measures can value locations of patches quite inconsistently. IIC-valued patch locations more in accordance with probabilistic measures than LCP, but showed deviations from the probabilistic measures at large threshold distances.

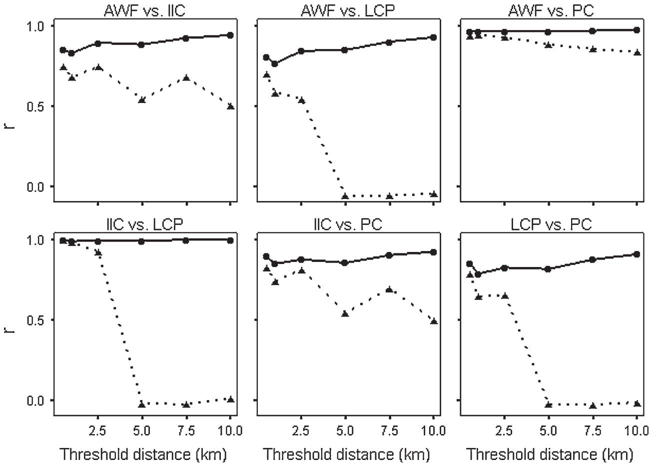
In the sparse network of herb-rich forests, the top-patches were even more inconsistently valued in prioritisations than all patches as a whole (Fig. 4). In this network, binary measures valued top patches more for their location than patches as a whole (results not shown), which seemed to increase differences in patch prioritisations derived from different measures. This is especially reflected in the increasing disparity between probabilistic and binary measures for the top patches. The difference was more pronounced at the intermediate threshold distances. In contrast, in the dense network of heath forests, different measures scored top patches with approximately the same consistency as all patches as a whole (Fig. 4). In that network, the binary measure IIC valued top-patch locations in accordance with probabilistic measures.

Fig. 3 Pairwise comparisons of patch prioritizations based on different measures (AWF, IIC, LCP, and PC) as a function of threshold distance for the networks of herb-rich and heath forests. Comparisons are based on Pearson's correlations coefficients calculated for the log-transformed patch importance values. Partial correlations show the correlations for the importance values after the effect of patch size is controlled for

(a) Herb-rich forests



(b) Heath forests

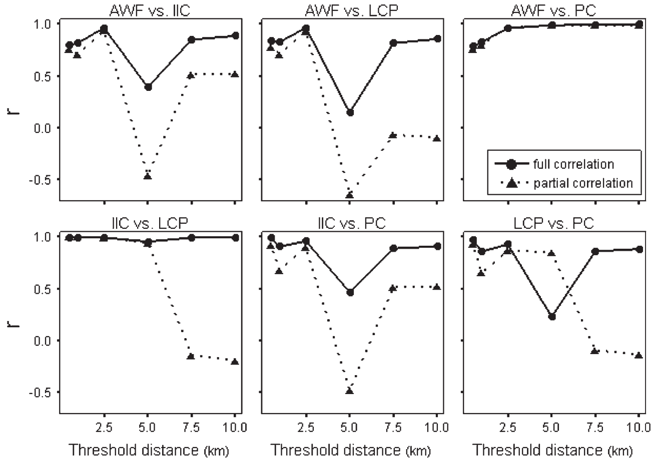


Discussion

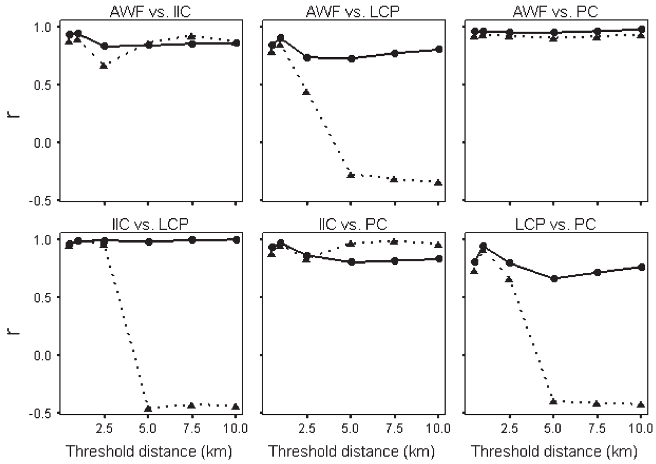
According to our conceptual and empirical comparisons, many connectivity measures do not react to changes in the amount and connection status of habitat in line with the connection concept defined. Measures that show ecologically problematic performance at the network level (graph diameter,

Fig. 4 Pairwise comparisons of patch prioritizations for 20 top patches based on different measures (AWF, IIC, LCP, and PC) as a function of threshold distance in the networks of herb-rich and heath forests

(a) Herb-rich forests



(b) Heath forests



correlation length, ECS) are not appropriate 'stand-alone' measures of landscape connectivity, but can nevertheless be used for descriptive purposes. Such descriptive information accompanied with a graphical presentation may give valuable insights into the characteristics of a network. On the other hand, connectivity measures responding in agreement with the connectivity concept at the network level differ greatly in terms of their connectivity model. Our empirical results give important insights of the implications of the connectivity model on the performance of measures at the landscape and patch level—not so evident when evaluated on conceptual grounds.

Because a small increase in dispersal ability (as well as a change in spatial pattern) may produce a sharp rise in connectivity, it has been suggested that connectivity possesses a transition range (e.g. With and Crist 1995; Keitt et al. 1997; and for the more general context of percolation theory, see Gardner et al. 1987; Stauffer 1987; With and King 1997). Transition range divides the species into two groups: those that perceive the landscape as disconnected and those that perceive it as connected. Our connectivity profiles (Fig. 1) indicate that the existence of the transition range is only evident for measures using a binary connection model. In contrast, measures that weigh interpatch distance by dispersal probability did not show a sudden shift in connectivity, but rather a gradual monotonic increase with increasing dispersal distance (as also reported by Saura and Rubio 2010). Empirical data would be needed to evaluate whether this transition range expressed by binary measures is also reflected in the realised ecological processes. If so, the binary measures may be used alongside with probabilistic measures as they are able to provide this additional insight to the interaction between landscape structure and species dispersal.

Our results show that connectivity evaluations of coherence measures are uninformative after the landscape consists of a single component (Fig. 1). Measures that view components as binary structures and do not account for their internal structure may be ecologically unrealistic. This especially holds for networks that show a high level of compartmentalisation. For example, if only one link (its length near the dispersal ability of organisms) is bridging two compartments, the dispersal of organisms may be restricted to within compartment rather than within component. The potential value of this one single link may be related to the process under focus and its associated time frame, because the probability of rare events increases as the time frame is extended. For processes that operate over long time periods (e.g. gene flow), weak connections among components may also be valuable, whereas for a process of a shorter time frame (e.g. population persistence in a fragmented landscape), the value of one weak connection is possibly only negligible. If it is necessary to correct this dependence on one link, an easy remedy would be to set a minimum number of links (greater than one) that should bridge a patch in a component, so that the component-based connectivity would lie on a stronger basis. Connectivity measures that are founded on the component approach (summarised in Table 1) are all confined to this possibly restrictive connectivity concept. Besides being widely used as a conceptual basis for connectivity measures, there has also been a growing interest in using components as landscape ecological units of analyses (e.g. Castellon and Sieving 2007; Minor et al. 2009).

Patch prioritisations based on different measures can produce inconsistent outcomes, which was evident also for the most valuable patches usually identified for conservation purposes (Figs. 3, 4). Our results show that the disparity between patch prioritisations derived from different measures may stem from two mechanisms: the measures place a different emphasis on patch size compared to patch location or they differ in how they value the locations of patches. When the results from graph analyses are interpreted, it is important to remember that the mere choice of a measure affects the relative importance of a patch area versus its topological properties. If the interpretation is biased in topological aspects, small patches may score too high relative to their ecological value. In our analysis, we interpreted the prioritisation value of a patch to be a combined outcome of its size and location in a network. The locational value of a patch can be further divided into two components; a patch can either be a flux donator/receiver in its own pairwise connections and/or a connecting element (i.e. a stepping stone) in the paths among other patches (Saura and Rubio 2010). This elaborated division may give additional insights into the behaviours of different measures in evaluating patches relative to their locations.

Networks that vary in patch density (depicted by the number of patches in a network per landscape area) provide testing grounds for the performance of connectivity measures. When there are more patches spread out in a landscape, their mean interpatch distance decreases, and the nature of their connections may also change (for example, the relative influence of indirect patch connections may increase in sparser networks). Dense networks may also exhibit a greater redundancy in interpatch connections than sparse networks, so that connections between pairs of patches are not necessarily only dependent on a single ‘route’. This redundancy is an important network property that should be understood when the results from node removal analyses are interpreted (as shown by Bodin and Saura 2010). Of course, a network with a small number of patches may have them all very locally gathered, and thus, from the point of view of configuration, show more characteristics of a dense network than a sparse one. This was not, however, the case in our networks, because the average nearest neighbour (NN) distance decreased with network density (herb-rich forests 924.97 m; spruce mires 441.51 m; heath forests 306.04 m). The average NN index (the ratio of the observed NN distance divided by the expected one based on random locations, calculated with ArcMap v. 9.2), which expresses the degree of patch clustering, was also at the same level in all three networks (herb-rich forests 0.64; spruce mire 0.59; heath forests 0.62). In our patch prioritisation results, the sparse network was more susceptible to inconsistent patch prioritisations than the dense one. We see that the conceptual differences between measures may become more evident in the patch prioritisations of sparse networks, but more specific mechanisms and generalisations would require prioritisation results to be analysed from a larger pool of networks. But regardless of network density, we recommend the comparison of prioritisation results from several measures to get a broader picture of the values and roles of patches in a network.

Keitt et al. (1997) suggested that patch configuration plays the most prominent role in landscape connectivity for species with intermediate dispersal ability, whereas for poor or good dispersers the patch value is more determinant. This is because poor dispersers are mostly confined to local patches irrespective of patch configuration, while good dispersers are not dispersal limited, and thus the patch size dominates the patch location in importance. Binary measures (IIC and LCP) and PC showed this pattern in their patch prioritisations, although PC valued configuration in a smoother manner than the binary measures (Fig. 2). However, AWF was not able to recognise the value of patch configuration at intermediate threshold distances (Fig. 2), suggesting that AWF cannot reliably value patch configuration from the perspective of dispersing individuals.

The relative influence of patch size (over configuration) on patch prioritisations was, at small threshold distances, greater in the sparsest network than in the two denser networks (Fig. 2). This indicates that individuals in sparse networks are confined to individual patches/isolated parts of a network. In such cases, the other patches are out of reach to individuals irrespective of their configuration, and habitat area as a currency of (intra- and interpatch) connectivity gets relative high values. Once again AWF deviated from the other measure of connectivity and did not detect this enhanced value of patch size in the sparsest network.

Our analysis showed that all connectivity measures react in their own characteristic ways (dictated by their connectivity model) to the removal of a patch. The mechanisms leading to evident differences in patch prioritisations among measures are intractable based on correlations alone, but would require additional investigation. The node removal analysis presents different kinds of challenges to and requirements on the connectivity measures than the measurement of landscape-level connectivity, because the evaluation of changes induced by patch removal brings about elements that are not exposed when working with intact networks. For the interpretation of the results from node removal analysis, it is important to understand how the given measure reacts to a connectivity change induced by patch removal. For example, the node removal analysis based on component-based measures cannot value a patch location unless the patch bridges otherwise isolated components. In the “non-split” cases, the prioritisation of a patch is only based on its contribution to habitat area. Measures based on the shortest paths (e.g. IIC and PC), on the other hand, may react unpredictably to patch loss because the measures also rate the possible new shortest paths formed in a network (Bodin and Saura 2010). Besides the patch prioritisations based on node removal analysis, there exist also alternative methods to evaluate the connectivity contributions of individual patches (e.g. Estrada and Bodin 2008). These alternatives, rooted in the concept of network centrality, operate on intact networks and have a very different philosophical background from the approach based on node removal analyses. As shown by Bodin and Saura (2010), these approaches may be used to complement each other in order to have a more comprehensive picture of patch value in a network.

Probabilistic connectivity measures track the enhancement in connectivity brought about by increasing dispersal ability of a species much further than do binary measures, which has implications for both landscape- and patch-level connectivity evaluations. Whether the realised connectivity tracks this enhancement captured by probabilistic measures may be related to the process under focus. As speculated by Bodin and Saura (2010), the binary connection model offered by IIC may be suitable for a process like gene exchange, which may be more reactive to a few realised transmission events (and so, to a mere existence of links rather than their strengths) than an actual transmission rate. The probabilistic connectivity models, on the other hand, may give a better basis for tracking quantities (e.g. dispersal flux) moving in a network.

Borgatti (2006) showed that the selection of optimal nodes may strongly depend on the goal. A different set of patches is selected when the purpose is to identify the patches whose removal would cause the greatest network fragmentation compared with a set that most efficiently promotes connections to other patches. This ultimate target of patch prioritisations is not usually explicitly stated or even considered in landscape ecological applications, although it should be inherent to the whole process of patch prioritisation. Graph-theoretic analyses can also be sensitive to the underlying assumptions regarding the dispersal behaviour of individuals. Connectivity at a landscape or patch level cannot be evaluated without a reference as to how individuals disperse in a network (e.g. Borgatti 2005; Urban et al. 2009). The most sophisticated graph-theoretic connectivity measures (IIC and PC) are based on the concept of the shortest paths. Insights provided by empirical connectivity data are needed to better understand which kinds of species and processes (and over which time frames) would possibly experience connectivity in terms of direct connections instead of paths. Moreover, dispersing individuals may not be restricted to the shortest paths, but alternative paths may also be important.

We would like to note that in many cases the actual *level* (quantitative variable) of connectivity is more important than judging whether a landscape is in the connected or disconnected *stage* (binary variable) for any given species. It is important to keep in mind that the total landscape area remains the same irrespective of the dispersal ability, and the level of connectivity is conditional on the total habitat area in a landscape. It is the network topology that determines how the connectivity changes with increasing dispersal ability. Conservation should offer both elements: habitat area as well as its appropriate configuration to allow efficient species dispersal. Even the sparsest of networks show levelling-off in connectivity with increasing dispersal ability, but there is no much use of this topological connectedness if the habitat area in a system is minimal. There have been recent attempts to minimise the total area of protected area network while maintaining the connectivity (e.g. Rothley and Rae 2005). We are afraid that connectivity is not yet understood thoroughly enough for this to be a safe approach. If no certainty exists that the connectivity measures capture the ecological process in focus, ecologically effective conservation should preferably address the question: How can we maximise connectivity for a given amount of habitat?

Scope for the future

The network robustness is an issue often neglected in the evaluation and design of habitat networks (see e.g. Matisziw and Murray 2008; McRae et al. 2008). Robust networks must contain nodes that are apparently redundant at the moment, in order to be resilient to future habitat loss or local extinctions. Robustness stems from the system redundancy, i.e. multiple pathways between pairs of patches. Matisziw and Murray (2008) deal with an important part of connectivity conservation by stating that ‘although ensuring a desired level of connectivity in a habitat network might be relatively easy to accomplish at the outset, long- term management of connectivity is much more complex’. Indeed, patch loss can impact connectivity in different ways and patch configurations differ in their robustness to the effects of patch loss. Most network measures do not hold information about network robustness, but only evaluate a network in one point of time. This is definitely an issue that needs further attention.

Graph-theoretic connectivity approaches most often view the connection between two patches as occurring via one link (the shortest/least-cost one), and additionally via one path (the shortest one) if indirect connections are also acknowledged. This is a problem that needs to be addressed, because in reality multiple connections certainly increase connectivity. This is supported by empirical evidence from a connectivity evaluation carried out in a circuit-theoretic framework showing that multiple connections between patches enhance gene flow (McRae and Beier 2007). Multiple connections can occur at two levels: between two patches that are in direct contact and at the larger network level, so that a patch pair is linked via more than one path. The former form of multiple connections has already received attention from several perspectives (e.g. Pinto and Keitt 2009; Urban et al. 2009; Vogt et al. 2009; Lookingbill et al. 2010), and if it can be compressed into one value depicting the isolation between two patches, it can be easily applied to any graph-theoretic connectivity measure. For example, circuit-theoretic resistance distance (a measure of isolation between pairs of patches) decreases with an increasing number of alternative connections between patches (McRae et al. 2008). The latter form of multiple connections, i.e. multiple paths among patches, has not yet been incorporated into graph-theoretic connectivity measures. Graph theory itself does not set limits, but provides potential for acknowledgement of more versatile connections among patches than have been acknowledged so far.

Acknowledgments This research was started as part of The Finnish Environmental Cluster Research Programme "Talousmetsät, pienialaiset suojelukohteet ja suojelualueet monimuotoisuuden suojelussa: Integroitu ekologinen vaikuttavuus ja kustannustehokkuus." The research was partially funded by Kone Foundation as a grant (to AL). We are also grateful to the Academy of Finland for funding (project #7115560, to MM). We thank the Forest and Park Service, the Regional Forestry Centre in Central Finland, the Regional Council, UPM Kymmene, and the Regional Environment Centre for providing data for our study. Pasi Reunanen helped us in network analyses conducted with the SELES program.

References

Andersson E, Bodin O (2009) Practical tool for landscape planning? An empirical investigation of network based models of habitat fragmentation. Ecography 32:123–132

Andren H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos 71:355–366

Bélisle M (2005) Measuring landscape connectivity: the challenge of behavioral landscape ecology. Ecology 86:1988–1995

Bender DJ, Contreras TA, Fahrig L (1998) Habitat loss and population decline: a meta-analysis of the patch size effect. Ecology 79:517–533

Bodin Ö, Saura S (2010) Ranking individual habitat patches as connectivity providers: integrating network analysis and patch removal experiments. Ecol Model 221:2393–2405

Borgatti SP (2005) Centrality and network flow. Soc Netw 27:55–71

Borgatti SP (2006) Identifying sets of key players in a social network. Comput Math Organ Theory 12:21–34

Bunn A, Urban D, Keitt T (2000) Landscape connectivity: a conservation application of graph theory. J Environ Manag 59:265–278

Castellon TD, Sieving KE (2007) Patch network criteria for dispersal-limited endemic birds of South American temperate rain forest. Ecol Appl 17:2152–2163

Estrada E, Bodin O (2008) Using network centrality measures to manage landscape connectivity. Ecol Appl 18:1810–1825

Fahrig L (1997) Relative effects of habitat loss and fragmentation on population extinction. J Wildl Manag 61:603–610

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34:487–515

Fahrig L, Merriam G (1994) Conservation of fragmented populations. Conserv Biol 8:50–59

Fall A, Fall J (2001) A domain-specific language for models of landscape dynamics. Ecol Model 141:1–18

Fall A, Fortin M, Manseau M, O'Brien D (2007) Spatial graphs: principles and applications for habitat connectivity. Ecosystems 10:448–461

Ferrari J, Lookingbill T (2009) Initial conditions and their effect on invasion velocity across heterogeneous landscapes. Biol Invasions 11:1247–1258

Ferrari J, Lookingbill T, Neel M (2007) Two measures of landscape-graph connectivity: assessment across gradients in area and configuration. Landscape Ecol 22:1315–1323

Fuller T, Munguia M, Mayfield M, Sanchez-Cordero V, Sarkar S (2006) Incorporating connectivity into conservation planning: a multi-criteria case study from central Mexico. Biol Conserv 133:131–142

Gardner R, Milne B, Turnei M, O'Neill R (1987) Neutral models for the analysis of broad-scale landscape pattern. Landscape Ecol 1:19–28

Goodwin BJ (2003) Is landscape connectivity a dependent or independent variable? Landscape Ecol 18:687–699

Gross JL, Yellen J (2006) Graph theory and its applications. Chapman & Hall/CRC Press, Boca Raton, FL

Hanski I (1999) Metapopulation ecology. Oxford University Press, Oxford

Jordán F, Báldi A, Orci K, Rácz I, Varga Z (2003) Characterizing the importance of habitat patches and corridors in maintaining the landscape connectivity of a *Pholidoptera transsylvanica* (Orthoptera) metapopulation. Landscape Ecol 18:83–92

Keitt T, Urban D, Milne B (1997) Detecting critical scales in fragmented landscapes. Conserv Ecol 1:4

Kindlmann P, Burel F (2008) Connectivity measures: a review. Landscape Ecol 23:879–890

Laita A, Mönkkönen M, Kotiaho JS (2010) Woodland key habitats evaluated as part of a functional reserve network. Biol Conserv 143:1212–1227

Lookingbill TR, Gardner RH, Ferrari JR, Keller CE (2010) Combining a dispersal model with network theory to assess habitat connectivity. Ecol Appl 20:427–441

Matisziw T, Murray A (2008) Connectivity change in habitat networks. Landscape Ecol 24:89–100

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. Proc Natl Acad Sci USA 104:19885–19890

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution and conservation. Ecology 89:2712–2724

Minor ES, Urban DL (2007a) Graph theory as a proxy for spatially explicit population models in conservation planning. Ecol Appl 17:1771–1782

Minor ES, Urban DL (2007b) A graph-theory framework for evaluating landscape connectivity and conservation planning. Conserv Biol 22:297–307

Minor ES, Tessel SM, Engelhardt KAM, Lookingbill TR (2009) The role of landscape connectivity in assembling exotic plant communities: a network analysis. Ecology 90:1802–1809

Moilanen A, Hanski I (2001) On the use of connectivity measures in spatial ecology. Oikos 95:147–151

Mönkkönen M, Reunanen P (1999) On critical thresholds in landscape connectivity: a management perspective. Oikos 84:302–305

O'Brien D, Manseau M, Fall A, Fortin M (2006) Testing the importance of spatial configuration of winter habitat for woodland caribou: an application of graph theory. Biol Conserv 130:70–83

Pascual-Hortal L, Saura S (2006) Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landscape Ecol 21:959–967

Pascual-Hortal L, Saura S (2008) Integrating landscape connectivity in broad-scale forest planning through a new graph-based habitat availability methodology: application to capercaillie (*Tetrao urogallus*) in Catalonia (NE Spain). Eur J For Res 127:23–31

Pinto N, Keitt T (2009) Beyond the least-cost path: evaluating corridor redundancy using a graph-theoretic approach. Landscape Ecol 24:253–266

Rothley KD, Rae C (2005) Working backwards to move forwards: graph-based connectivity metrics for reserve network selection. Environ Model Assess 10:107–113

Saura S, Pascual-Hortal L (2007) A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landscape Urban Plan 83:91–103

Saura S, Rubio L (2010) A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. Ecography 33:523–537

Saura S, Torné J (2009) Conefor Sensinode 2.2: a software package for quantifying the importance of habitat patches for landscape connectivity. Environ Model Softw 24: 135–139

Stauffer D (1987) Introduction to percolation theory. Taylor & Francis, London

Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. Oikos 68:571–573

Tischendorf L, Fahrig L (2000a) On the usage and measurement of landscape connectivity. Oikos 90:7–19

Tischendorf L, Fahrig L (2000b) How should we measure landscape connectivity? Landscape Ecol 15:633–641

Tischendorf L, Fahrig L (2001) On the use of connectivity measures in spatial ecology. A reply. Oikos 95:152–155

Urban D, Keitt T (2001) Landscape connectivity: a graph-theoretic perspective. Ecology 82:1205–1218

Urban DL, Minor ES, Treml EA, Schick RS (2009) Graph models of habitat mosaics. Ecol Lett 12:260–273

van Langevelde F (2000) Scale of habitat connectivity and colonization in fragmented nuthatch populations. Ecography 23:614–622

Vogt P, Ferrari JR, Lookingbill TR, Gardner RH, Riitters KH, Ostapowicz K (2009) Mapping functional connectivity. Ecol Indic 9:64–71

Wiegand T, Revilla E, Moloney KA (2005) Effects of habitat loss and fragmentation on population dynamics. Conserv Biol 19:108–121

With KA, Crist TO (1995) Critical thresholds in species' responses to landscape structure. Ecology 76:2446–2459

With KA, King AW (1997) The use and misuse of neutral landscape models in ecology. Oikos 79:219–229

III

SPACE, SCALES AND IMPERFECT DETECTION: MODELING THE HABITAT ASSOCIATION OF THE SIBERIAN JAY (*PERISOREUS INFAUSTUS)* IN BOREAL FOREST LANDSCAPES

by

Anne Laita & Mikko Mönkkönen

Submitted manuscript

SPACE, SCALES AND IMPERFECT DETECTION –MODELING OF THE HABITAT ASSOCIATION OF THE SIBERIAN JAY (*PERISOREUS INFAUSTUS)* IN BOREAL FOREST LANDSCAPES

Anne Laita & Mikko Mönkkönen

Department of Biological and Environmental Science, PO Box 35 (Survontie 9),
40014 University of Jyväskylä, Finland

ABSTRACT

Species distribution models that combine sound ecological theory and reliable modeling approaches can be very useful tools in devising efficient conservation programs for species threatened because of habitat loss. The Siberian Jay (*Perisoreus infaustus*) is a representative of the resident species of northern taiga forests, whose declines have been attributed to the adverse, multiple-scale changes in forest characteristics brought about by intensive modern forestry. Here, we modeled the territory and landscape level habitat associations of the species in two areas in southern Finland (central Finland and Karelia) within a Bayesian, hierarchical model setting. In the modeling, we accounted for imperfect detection inherent in the occurrence data, as well as spatial autocorrelation in the response with the Gaussian conditional autoregressive (CAR) approach. Multi-scale models, incorporating territory and landscape level variables, outperformed single-scale territory level models, which highlight the role of landscape context in shaping the distribution pattern of the species. For territory level models, the spatial autocorrelation of the missing landscape level variables was incorporated to the spatially structured error term, improving the model fit and performance for both areas. In central Finland characterized by a patchy species distribution pattern outside of the continuous distribution range, the species seemed to set stricter demands on its habitat than in Karelia that is connected with the continuous range in the north. In central Finland, the CAR modeling coupled with multi-scale approach lead to further enhancement in model fit and differentiation ability, which suggests the role of intrinsic factors shaping the species distribution pattern. The successful conservation of the species requires forest planning at the sufficiently large scale, and special emphasis should be given to the structural characteristics of forest stands.

INTRODUCTION

Modern forestry has altered the dynamics and structure of the Fennoscandian forest ecosystems, albeit the forest land area has remained unchanged. Intensively managed forests are characterized by even-aged, single-storied stands, altered tree species composition, low amount of dead wood, the lack of old deciduous trees and continuity (Esseen et al. 1997, Östlund et al. 1997). At the landscape level the area of old-growth and mature forests has decreased, and become separated by younger stands and clear-cut areas (Löfman and Kouki 2001). Forestry has affected adversely the state of forest biodiversity, for example, 36% of threatened species in Finland are forest dwellers (Rassi et al. 2010). Among the species threatened by forestry are resident forest species, specialized in later forest successional stages that inhabit the northern boreal forests in Fennoscandia (Helle and Järvinen 1986, Virkkala 1987, Edenius and Elmberg 1996).

Species distribution models (SDMs) aim at quantifying the relationship between species occupancy and the environmental factors that impose control on it (Guisan and Zimmermann 2000). Species distribution models can be used to gain new insights into species habitat associations, but also their predictive capabilities can be widely utilized, for example in conservation planning and targeting of field inventories. Indeed, species distribution models that combine sound ecological theory and knowledge with the modeling approaches and methods giving justice to the properties of data and the process under focus, can be very useful tools in devising efficient conservation programs for species threatened because of habitat loss.

Habitat selection of organisms is a hierarchical process that occurs on several spatial scales (Wiens 1989, Levin 1992, Cushman and McGarigal 2002). The environmental variables that control the distribution of a species, and their relative importance, change with the spatial scale (see e.g., Buler et al. 2007). Thus, the successful modeling of species' distribution pattern usually calls for incorporating multiple spatial scales (see e.g., Naugle et al. 2000, Graf et al. 2005, Betts et al. 2006). Moreover, forestry affects forest characteristics and structure across the different spatial scales, from stand level to broad landscape level. The realized distribution pattern of a forest dependent species is a complex interplay of hierarchically assembled response of the species to its hierarchically modified environment.

Absence of a species at a site is more difficult to verify than its presence. Occupancy data for many species are burdened by imperfect detection so that a detected absence may not equate to true absence but undetected presence (see e.g., MacKenzie et al. 2002, MacKenzie 2006). To make inference of the actual state variable (species occurrence) and its relationship with environmental covariates, it is important to explicitly account for the imperfect detection inherent in occupancy data (e.g., Royle et al. 2007). Ignorance of imperfect detection in species-habitat modeling has been shown to bias the parameter estimates for habitat covariates and reduce their precisions (Tyre et al. 2003, Gu and Swihart 2004).

The distribution pattern of a species is not a straightforward product of environmental variables, but species' intrinsic properties (e.g., social behavior and dispersal) may also shape the realized pattern (Legendre 1993, Fortin and Dale 2005, Bahn et al. 2006). Usually this creates a more aggregated and autocorrelated distribution pattern that would be predicted based on the environmental variables alone. Along with intrinsic factors, also missing (or incorrectly specified) environmental (i.e., extrinsic) factors lead to autocorrelated residuals (Legendre 1993, Fortin and Dale 2005), and the two sources of autocorrelation are difficult to disentangle (see Teeffelen and Ovaskainen 2007). In addition to the possible predictive benefits achieved when the autocorrelated patterning is accounted for in distribution modeling (e.g., Miller et al. 2007), spatial approach also allows for making sound inference about the relationship between species' distribution and environmental variables controlling it. If a model violates the assumption of independent errors (as usually non-spatial models do), it tend to produce biased parameter estimates and overestimate the effects of environmental covariates on controlling species distributions (e.g., Dormann 2007).

Bayesian hierarchical models provide a flexible framework to model complicated processes, which can be viewed and modeled as an outcome of several conditional processes, specified via their own submodels (Wikle 2003). For example, in the case of imperfect detection the observational data do not allow direct inferences to be made on the actual process variable (i.e. species occurrence). Within a hierarchical framework, the process of interest can be conditioned on the observational data, and both (data and process) represented by their own submodels (Royle and Dorazio 2008). Bayesian model implementations becomes especially appealing or represent the only possible option for analysis when the modeling involves hierarchical formulation coupled with spatially explicit approach (see e.g., Latimer et al. 2006, Royle et al. 2007, Wilson et al. 2010). Moreover, the quantification of uncertainty associated with parameter estimations is also an evident advantage that comes along with Bayesian approach (see e.g., Cressie et al. 2009, Link and Barker 2010). The spatially explicit, hierarchical Bayesian approach offers an obvious benefit for species distribution modeling as it allows, along with predictive benefits, safe inferences to be made on the role of environmental covariates shaping the species distribution pattern. Thus, such modeling would provide a sound basis for species conservation programs.

The Siberian Jay (*Perisoreus infaustus*) is a sedentary passerine species that inhabits the boreal regions in Eurasia. The Siberian Jay has a complex social behavior: it lives in territorial social groups in which the breeding pair may be accompanied with retained offspring and/or non-kin immigrants (see e.g., Ekman et al. 1999, Ekman et al. 2002, Griesser et al. 2006). The number of Siberian Jay individuals in Finland declined to one third during 1940-1970, after which populations in southern Finland, where the species lives at the southern edge of its distribution range, have shown the further decline (Väisänen et al. 1998). Consequently, the distribution of the species has become highly patchy. The population decline has been attributed to the effects of intensive logging and forest management (e.g., Helle and Jarvinen 1986).

In southern Finland, the Siberian Jay has been suggested to exhibit clearly different and more stringent habitat requirements than in the north (e.g. von Haartman et al. 1963-72). The Siberian Jay is traditionally considered being associated with old-growth coniferous forests /swamp forests (Väisänen et al. 1998), but the species also occurs in intensively managed forests (Lillandt 2000). In addition to high quality breeding areas, rather large territories (up to 5 km^2) of the species typically comprise many kinds of forest stands and land cover types (also open areas, young stands and mires) (Lillandt 2000). As only a rather small proportion (6.8 per cent on average) of the area known to be occupied by the Siberian Jay is located in protected areas in Southern Finland (Pihlajaniemi 2009), the viability of the Siberian Jay populations relies strongly on the forest practices carried out on commercial forests. However, to find a safe balance between forestry practices and species habitat requirements, more knowledge is needed on the key determinants of its habitat selection and breeding success. Especially, little is known about the habitat associations of the species at the landscape level.

Siberian Jays show a high level of site fidelity and do not abandon their territories easily. Dispersal of the species is mainly restricted to short distances, usually near their natal territory (Lillandt 2000, Uimaniemi et al. 2000). The documented average dispersal distances in Southern Finland range from ca. 3 km (Lillandt 1993) to ca. 5 km (Gienapp and Merilä 2011). In addition to the dominating short distance dispersals, also long distance, inter-population dispersals have been confirmed (Lillandt 2000). Because of the complex social behavior of the species and its restricted dispersal it is to be expected that, besides the environmental determinants, the distribution pattern of the species may also be shaped by intrinsic autocorrelation.

We use here data from experimental feeding sites. The detection histories for intensively monitored occupied sites indicate that the probability of seeing a Jay in any given visit to a site is rather low. As the feeding site data obviously suffers from imperfect detection, we wanted to model the detection process at the feeding sites simultaneously with the occurrence modeling. In this study we address the following questions: 1) What are the territory and landscape level variables explaining the recent distribution pattern of the species in our study areas? 2) Is the distribution pattern of the species determined by territory level environmental variables, or is it also controlled by the forest characteristics and composition at the landscape level? 3) Does the accounting for spatial autocorrelation in the response provide a better understanding of the distribution pattern of the species than a non-spatial approach? 4) How consistent are the habitat associations of the species in different parts of southern Finland? Our research provides new information on the habitat-relationship of the species in managed forest landscapes, so that the key elements of the Siberian Jay forests could be better identified and acknowledged (at a proper spatial scale) in managed forests.

METHODS

Study areas

Study areas were located in central Finland and Karelia (Fig. 1). The study areas are separated by ca. 100 km in east-west direction, and the southern border of the study area of central Finland lies ca. 70 km southwards from that of Karelia. The study area in central Finland encompassed *c*. 21 500 km^2, of which 73 % was covered by forested land (of which forests 77 % and scrubland 23 %), 12 % by waterbodies, 11 % by built-up and agricultural areas, 3 % by open mires and < 1% by barren land. The study area in Karelia encompassed *c*. 18 500 km^2, of which 75 % was covered by forested land (of which forests 72 % and scrubland 28 %), 13 % by waterbodies, 6 % by open mires, 5 % by built-up and agricultural areas and < 1 % by barren land. Central Finland has a longer and more intensive history of land use than Karelia. In both study areas the dominant tree species are Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), and birch (*Petula pendula* and *P. pubescens*). Spruce is a dominant tree species in 26 % and 19% of the forested land in the study area of central Finland and Karelia, respectively.

Although the known presences of the Siberian Jay are spread throughout central Finland, its occurrence in the study area is very patchy. The study area of Karelia has higher Siberian Jay abundance than central Finland, and near the Eastern border this area represents a continuous distribution range that extends from the foothold areas of the Siberian Jay in Northern Finland (Pihlajaniemi 2006).

Forest data

We used the multi-source Natural Forest Inventory (NFI9) data produced by The Finnish Forest Research Institute as land use and forest characteristic data. This dataset is a combined product of field measurements, satellite image data and other digital data sources (elevation models and land use maps; Tomppo 1993). The resolution of the data is 25 m * 25 m. In our research we used layers for timber volume (m^3/ha) by tree species (spruce, pine, birch and other deciduous trees), basal area (m^2/ha) and forest age. We also used a layer that distinguished heath forests, spruce mire forests, forested mires/bogs and open fens from each other. In addition to the NFI –data, we used National CORINE database of

Finland to distinguish human dominated areas and a digital elevation model (DEM) to calculate elevation and topographic variation of terrain, both having a resolution of 25 m*25 m.

Raster datasets were processed in two steps. First, moving window analyses for the extraction of covariates (for all the covariates represented in Table 1) were conducted with the original data resolution (25 m * 25 m). We used a circular window with radius of either 750 m (territory level) or 5 km (landscape level). The circular window of radius 750 m corresponds in area (1.8 km^2) to the size of the Siberian Jay territory in Southern Finland, 1-5 km^2 (Lillandt 2000). The scale of landscape level analysis (5 km) is determined on the basis of the documented average dispersal distances. Second, the layers from the moving window analyses were resampled (using average criteria) to the resolution of 1.5 km * 1.5 km (2.25 km^2). We wanted to conduct the modeling with the resolution of an average territory size, because this provides an ecologically meaningful interpretation for autocorrelation – the possible dependence between territories and their occupancy status. However, the two-step extraction of covariates allowed us to exploit the information in the forest data more efficiently than straightly resampling it to the coarse resolution. We peeled our study areas by 5 km wide buffer strips (corresponding to the scale of landscape analysis) to have complete data coverage also near the borders.

Occurrence data of the Siberian Jay

We utilized two kinds of species occurrence data: one originating from repeated feeding site surveys and the other from the sighting record databases governed by local ornithological societies. The records of the both datasets are dated to the years 2005-2008. We assumed constant occupancy status of the territories during this study period.

The repeated feeding site surveys were conducted by ornithologists and two fieldworkers of Metsähallitus (one in central Finland and another in Karelia). The feeding site survey is not based on systematic sampling design, which poses challenges to the efficient use of the data set.

The feeding site survey relies on the hoarding behavior of the Siberian Jays in autumn. Feeding sites are provided with a bag made of wire netting, filled with lard. Within jay territories, feeding sites attract individuals and facilitate the detection of individuals. We coded the feeding site data as a total number of visits to a site and a number of visits with jay detections (either sighting or hearing). If there were several feeding sites for one grid cell, the data entries of individual feeding sites were combined. If the total number of visits was not reported for a site, we used a median value. The median was calculated primarily based on the data (i.e. total number of visits per site) from the other feeding sites of a given observer. If that data was unavailable, the median was calculated from the area-survey specific data. There was a great variation in the number of visits made for each site, ranging from 1-54 and 2-6 for central Finland and Karelia, respectively.

Sighting record databases comprised a collection of sightings reported by ornithologists and public. From the records we removed the ones with inaccurate locations (when either the precision of coordinates was inaccurate, or location was reported to be unreliable).

The two sources of species occurrence data were considered to differ in terms of detection process. Detection probability was modeled only for repeated feeding site surveys, but the sighting records were considered to represent perfect detection ($p=1$). For feeding site surveys, our dataset comprised 402 grid cells (of the resolution of 1.5 km* 1.5 km) for central Finland and 145 for Karelia. For sighting records, there were 99 grid cells for central Finland and 46 for Karelia. Altogether, there were 103 and 47 grid cells with detected occupancy in central Finland and Karelia, respectively.

Environmental covariates and their ecological basis

The territory ($n=14$) and landscape ($n=5$) level covariates we used in modeling are summarized in Table 1. The choice of the environmental covariates was based on the existing knowledge on the ecology of the Siberian Jay. Here, we provide a brief description of the ecological basis of the covariates as well as the details how the covariates were calculated.

In southern Finland, the Siberian Jay has traditionally been considered to be associated with old, continuous, spruce-dominated forests (e.g., Väisänen et al. 1998), but it also occurs in commercially managed forests. Spruce offers concealment from avian predators as well as diverse feeding and nesting opportunities (Edenius and Meyer 2002). On the other hand, the Siberia Jay has been shown to avoid open spruce-dominated forests with large diameter trees (Lillandt 2000). Especially in Southern Finland, the species further is considered to be associated with spruce mires (e.g., Väisänen et al. 1998), which possibly provide suitable, multi-layered forest structure (rich in low spruces) for nesting and hoarding. Young forests are assumed to play a negative role for a species as they represent rather open forest areas and forested areas unsuitable for nesting (i.e., regeneration sites, seedling stands and young thinning stands).

We used three cover variables. First, forest cover (total volume >100 m^3/ha) tested if overall availability of forests is sufficient in explaining the distribution of the species. Second, cover of spruce dominated forests (total volume >100 m^3/ha) was entered to test for the presumed association with spruce, and, third, cover of spruce dominated matured forests (total volume > 170 m^3/ha) to test for association with older spruce forests. These three variables were highly correlated, and they were not allowed to enter simultaneously to models. The growing stock volumes of 100 m^3/ha and 170 m^3/ha corresponds to those of an average young thinning stand and advanced thinning stand, respectively

(Anonymous, 2009). A pixel was defined to be spruce-dominated if the growing stock volume of spruce exceeded that of pine or deciduous trees. As additional forest cover variables we calculated the extent of spruce mires (total volume > $100m^3/ha$) within territory and the extent of young forests (< 40 years) at the territory and landscape level.

Nesting success of the Siberian Jay has been found to be enhanced in the vicinity of "soft" forest edges (Sklepkovych 1997). According to Skepkovych (1997), forest –mire edges may be rich in food because sun exposure at the transitional areas increases insect activity and provides areas of earlier snowmelt. Forest structure in these transitional areas may also be favorable to the species. We calculated the total length of soft forest edge between forests with total timber volume > $100m^3/ha$ and mires (mires and fens). The raster dataset was generalized by replacing all cell groups ≤ 6 cells (with a four neighbor criteria) with the values of the nearest neighbors. In addition to the length of soft edges, we also calculated the coverage of mires and fens, because open or sparsely forested areas has also been found to increase adult mortality through enhanced predation risk (Nystrand et al. 2010)

The elevation above the sea level has been found to be an important broad scale determinant of Siberian Jay occurrences in North Finland where the jays had been found to favor areas with elevation above the sea level greater than 235 m (Louhia 2001). There is no apparent causal mechanism explaining this association, as the elevation may, in addition to its independent effect, also be intertwined with many forest cover variables. We prioritized thus the cover related variables over elevation and terrain variables, which were entered at the later stage in models (Fig 2.). The variation in elevation (i.e., topographic variation) within a territory may also bring about environmental heterogeneity, which may have a positive association with the occurrence of the Siberian Jay.

The Siberian Jay has been shown to prefer areas located further away from human settlement (Ekman et al. 2001, Nystrand et al. 2010). Although the influence of human settlement is most certainly linked with cover and terrain variables as well, it could also have an independent role in controlling the species occurrence. The negative effect of human settlement may be mediated by other corvid species, which thrive near human settlement and whose predation is the primary cause of the Siberian Jay nest failure (Eggers et al. 2005). We calculated the influence of human settlement as the proportion of human-dominated land use (including built-up and agricultural areas) of the total territory/landscape area.

Forest openness within territories increases the exposure of jays and their nests to the visually hunting predators (Eggers et al. 2005, Griesser et al. 2007, Nystrand et al. 2010). We calculated tree density for spruce-dominated forests (> 20 years) as the ratio of basal area (m^2/ha) to the average height of a forest stand (m). We used the following formula to first calculate the average stock height: total timber volume (m^3/ha) = shape index*basal area (m^2/ha)*average stock height (Auvinen 1997), using a value of 0.5 for a shape index. The density layer (basal area/stock height) was smoothed by moving window analysis with median criteria. The smoothed density layer was then resampled to a final resolution using either average (Densityave) or maximum (Densitymax) criteria. Smoothing was necessary, because timber volume and basal area layers are not in complete match at the pixel level, leading to some extreme density values. The smoothed layer gives rather conservative information on density, as it also wipes out information on the true extreme density values (which could be ecologically more important than the average ones).

There is a SW to NE gradient in density of the Siberian Jay in Finland. The species occurs as patchy populations in southern Finland, and the southern edge of the range boundary has shifted towards north during the last decades (Kemppainen and Kemppainen 1991, Lillandt 2009). Further north, densities tend to increase towards east due to the supply of individuals from the Russian Karelian continuous forests (Heikkinen et al. 2000, Kouki and Väänänen 2000). To control for the gradient, we used x- and y-coordinates (expresessed as the distance from the western and southern border of the study area, respectively) as variables.

The decline of the Siberian Jay in Finland has been greater than predicted on the basis of the decrease in the area of old-growth forests alone (Järvinen et al. 1977, Helle and Jarvinen 1986). This indicates that, along with habitat loss, forest fragmentation may also be involved in the declining population trends. We used average patch size of spruce-dominated forests (> 170 m^3/ha) at the landscape level, included to the models only simultaneously with the corresponding cover variable (i.e., cover of spruce dominated forests at the landscape level).

For the effect of mire areas (variables Edge and Mire) as well as of young forests (Young), we also entered quadratic terms to allow non-linear patterns to be modelled. For Edge and Mire this was *a priori* decision, and for the variable Young, a *posteriori* decision.

We used FRAGSTATS v. 3.3 (McGarical et al. 2002) to calculate the total length of edges between forests and mires/fens as well as the area-weighted mean patch size and GRASS v. 6.4 (GRASS Development Team 2010) for calculations of all other variable layers.

Modeling procedure

As the feeding sites were not randomly located in the two landscapes, it was to be expected that they do not provide a representation of an average landscape nor do they possibly capture the full range of variation for the environmental covariates. We checked this by creating 500 random points in both areas (located on forested land with growing stock volume >100 m^3/ha) and comparing their distributions for territory and landscape level covariates with corresponding distributions of the feeding sites and sighting records. The comparisons of univariate distributions showed that the survey efforts were biased towards high elevation- low human impact sites (results not shown). As we wanted to correct for this survey bias, we opted for two-step modeling procedure. The first step was used to create pseudo-absence points that were fed as absence data to the actual modeling efforts of the second step (see Zaniewski et al. 2002, Engler et al. 2004, Chefaoui and Lobo 2008 for the description and applications of this methodology).

The pseudo-absence points were produced by using environmental niche factor analysis, ENFA. ENFA compares, in the multidimensional space formed by environmental covariates, the distribution of the known presence points with the reference distribution of the whole landscape (i.e. global distribution) (Hirzel et al. 2002). ENFA extracts differences in those two distributions, in terms of their mean positions and variances, to uncorrelated factors. The first factor, marginality factor, maximizes the differentiation of the species niche from the global distribution, while the subsequent specialization factors maximize the variance ratios of the two distributions. The extracted factors allow the computation of the habitat suitability scores for all the cells of the study area. We used eight box-cox transformed territory level variables (CoverT, Edge, Elev, HumanT, SmireT, Spruce170_T, YoungT, Densitymax) for factor extraction, and retained four factors for both areas for the calculation of habitat suitability scores. Habitat suitability scores, ranging from 0 to 100, were calculated with median algorithm, with adjustment for global frequency and extreme optimum (see Braunisch et al. 2008). The choice of the algorithm and its parameters was based on the inspections of the ten-fold cross-validations results (predicted/expected curves and Boyce indice) (see Hirzel et al. 2006). Pseudo-absence points were randomly located to those forested cells that have habitat suitability (HS) scores ≤ 10 and growing stock volume >100m^3/ha. For both areas, the number of pseudo-absence points equaled the number of the known presence sites. Cells with HS scores ≤ 10 covered more than half of the landscape area in both two landscapes, and included only c. 10% of the presence records. The ENFA results for both areas are represented in Appendix I. As for the sighting records, the detection probability for the pseudo-absence points equaled one in the modeling.

We constructed two separate sets of candidate models for both areas; one comprehending only territory level covariates and another combining territory and landscape level covariates. Figure 2 provides a schematic chart of the order in which we allowed the covariates to enter the models. For landscape models, we were interested in the additional explanatory power of landscape level variables, after the territory level variables have been accounted for. We also tested whether the landscape level analogue of a territory level variable had a better explanatory power than the territory variable itself. Because the number of candidate models was high, we decided to make model comparisons in a computationally “light” manner (as opposed to computer intensive Bayesian methods). We used program Presence (v. 3.0) (Hines 2006) to fit a set of candidate non-spatial logistic regression models, where we modeled occurrence probability as a function of environmental covariates, while allowing for imperfect detection for the feeding site records (MacKenzie 2006). The candidate models were compared based on the AIC$_c$ values (Burnham et al. 2002), with effective sample size corresponding to the number of sampling units.

For the best models based on AIC $_c$ values, we fitted Bayesian logistic regression counterparts as well as their spatial realizations. As the best models, we recognized the models whose AIC$_c$ differences (Δ_i; AIC$_{ci}$ minus AIC$_c$(*Min*)) were < 4. According to the interpretation by Burnham and Anderson (2002), models with $\Delta_i > 4$ have considerably less support than the model with minimum AIC value.

Bayesian analyses were conducted with OpenBugs v. 3.0 (Thomas et al. 2006). We employed a Gaussian conditional autoregressive (CAR) modelling approach to account for the spatial autocorrelation in the response. CAR models incorporate autocorrelation in a spatial error term (Besag 1974). For the CAR applications in the context of species distribution modeling, see e.g., Farnsworth et al. 2006, Latimer et al. 2006, Carroll et al. 2010 and Ishihama et al. 2010. In CAR model, probability of species presence (psi_i) depends, along with effect of environmental covariates, on a site-specific spatially-structured random effect. This random effect is influenced by the values of the neighboring sites.

CAR model has the following formula:

$$logit(psi_i) = \alpha + \sum_k \beta_k \, x_{ki} + \rho_i + \varepsilon_i,$$

where α is the intercept; β_k is the coefficient for the explanatory variable x_k; ρ_i is the spatial random effect term for the site i ; and ε_i is the spatially non-structured random error. The spatial random effect term of a site i (ρ_i) is dependent on the ρ terms of the neighboring sites. We defined the neighborhood of cell i to encompass its eight adjacent cells. The following conditional relationship is assumed for the distribution of ρ_i:

$$\rho_i \mid \rho_{i+} \approx \text{Normal}\left(\frac{\sum_{j\in\delta i} \rho_j}{n_{i+}}, \frac{\sigma_\rho^2}{n_{i+}}\right),$$

where $i+$ stands for a set of neighbors of cell i; cell j belongs to the neighborhood of cell i (denoted by δi); n_{i+} denotes the number of cells which are neighbours of cell i; and σ_ρ^2 is the variance for all cells.

We used uninformative priors for model parameters (uniformly distributed unif(0,1) prior for detection probability, normally distributed dnorm (0, 1.0E-6) priors for the intercept and β parameters, and the prior specification used by Latimer et al. 2006, an inverse gamma prior, for the conditional variance σ_ρ^2 in a spatial error term). We run the models for computational reasons with one MCMC chain (but for randomly selected model runs we used two chains to control for the convergence). We used a burn-in period of 20 000 -40 000 iterations, followed by 20 000 iterations for parameter estimation.

Evaluation of the model performance

We evaluated the fit and performance of the final Bayesian models based on their 1) deviance values; 2) Minimal Predicted Areas; and 3) cartographic output of the predictions. Because many of the methods that are usually used for the evaluation of model performance require absolute knowledge on the site occupancy (i.e., area under curve (AUC), true skill statistics (TSS), mean-squared predictive error (MSPE)), they turn infeasible for data with imperfect detection. Minimal Predicted Area (MPA), first introduced by Engler et al. (2004), stand for the extent of the predicted surface obtained by only considering those grid cells whose probability value exceed the threshold value, set to exclude x % (in our case, 10 %) of the known presences with smallest predicted probabilities. MPA value thus represents the proportion (%) of the area (of the total study area) exceeding the threshold predicted probability value. The leading idea behind MPA is that a good model with a high discriminatory capacity should predict a potential area that is as small as possible while still comprising the most of the species occurrences. The logic of the MPA is closely related to that of the AUC, while the measure also allows for visual map interpretation (Latimer et al. 2006).

In addition, we constructed the Moran's I correlograms for the model residuals to check for the spatial autocorrelation left in different distance classes (bin width of 2.5 km approximately equaling to distance between a cell and its cardinal neigbours (2.25 km). Because the true occupancy state of the grid cell with imperfect detection was unknown, we only qualified those imperfectly detected grid cells whose predicted probability of occupancy was < 0.4 based on the predictions of the best-performing multi-scale CAR model of the area (L1 for central Finland and L2 for Karelia).

RESULTS

As a result of model selection based on AIC_c values, we ended up with four and three best models both at the territory level (Appendix II) and at the multi-scale level (Appendix III) for central Finland and Karelia, respectively. The variables in the models were not strongly correlated ($r < 0.5$). The parameter values for their spatially explicit CAR model counterparts are represented in Table 2 (for territory level models) and Table 3 (for multi-scale models). Parameter estimates and their standard deviations remained largely unchanged in the spatial models, compared to the values of the corresponding non-spatial models. For both areas, the probability of detecting at least one Siberian Jay individual (given the territory occupancy) in one visit was rather low, *ca.* 0.10, across all the candidate models. For the reference, the null-model AIC_c values (including the terms only for the intercept and the detection probability) were 1032.8 ($\Delta_{AICc}=82.3$) and 282.8 ($\Delta_{AICc}=43.4$) for central Finland and Karelia, respectively.

Model selection procedure resulted in rather different set of variables being selected to the best-screened models for central Finland than for Karelia (Appendices II and III, Tables 2 and 3). The best-supported territory level models for central Finland suggested positive effects of the length of edges between forests and mires/fens, the cover of young forests, and the cover of spruce mires (Fig. 3a), and a negative effect of the cover of human dominated areas on the probability of Siberian Jay occurrence . Although the variables for the length of edges and human-dominated areas were not strongly correlated (r=-0.41), they seemed to carry overlapping information on the species habitat association resulting in two sets of models for the species, one with edge length (T1, T2) and another with the cover of human dominated areas (T3, T4) included (Appendix II, Table 2)

For Karelia, the best-supported territory level models suggested positive effects of elevation and y-coordinate (northing) on the probability of occurrence, which together comprised the model with the minimum AIC_c value (T1). (Appendix II, Table 2). However, the inclusion of the variables for the cover of spruce-dominated forests (total volume >170m^3/ha) (positive, Fig. 3b), the cover of mires (unimodal relationship) and the cover of young forests (positive) improved the MPA performance of the models (T2). The elevation variable dominated the models for Karelia, because terms of the other variables turned insignificant when occurring together with it (T2 vs. T3). Although the AIC_c difference for the model without the elevation was huge (Δ_i=15), its MPA performance was highest among the territory level models for Karelia. The inclusion of forest density variable (positive effect) to the models was slightly supported by the AIC_c values in both areas, but the models with forest density variable did not lead to improved performance (based on MPA), nor were their terms significant in all the models (Appendix II, Table 2).

For central Finland, the multi-scale models with the highest support indicated a positive landscape level effect of forest cover (total volume > 100 m^3/ha) (Fig. 3c) and a negative landscape level effect of human-dominated areas. The landscape level variable for human-dominated areas clearly dominated its territory level counterpart in explanatory power. As with the territory level models, the variables for edge length and cover of human-dominated areas at the landscape level (r=-0.39) did not either occur in a same model, or occured together both with insignificant terms (L2) (Table 3).

For Karelia, the landscape level variables for the cover of young forests (positive effect), human-dominated areas (negative effect, Fig. 3d) and mires (and its quadratic term; although insignificant together with the elevation variable in the models) were included among the best-screened multi-scale models (Table 3). For both areas, multi-scale models outperformed territory-level models based on AIC_c values and MPA-performance (Appendices II and III), indicating the importance of landscape context in determining the distribution pattern. The superiority of multi-scale models was especially clear for the non-spatial models.

The CAR approach resulted in a remarkable improvement in model fit and performance compared to the non-spatial approach. The benefits of the spatial approach were more pronounced for the territory level than for the multi-scale models in both areas (cf. Appendix II with Table 2 and Appendix III with Table 3). For central Finland, however, the spatial approach clearly outperformed the non-spatial approach also for multi-scale models (Appendix III and Table 3). For example, in multi-scale models for central Finland, the MPAs dropped from the level of *ca.* 39 % to *ca.* 20% of the study area when switched from the non-spatial to the spatial approach. The associated deviance values also decreased from the level of *ca.* 950 to *ca.* 870. The predictive distribution patterns based on the CAR-approach were much more aggregated compared to the prediction produced by non-spatial models (Fig. 4). Instead, the fit and MPA performance of the multi-scale non-spatial and CAR- models were approximately at the same level for Karelia (Appendix III and Table 3), producing similar predictions for distribution pattern (Fig. 5., only showing the predictive surface for the multi-scale L2 model, almost identical to that produced by the non-spatial counterpart). Generally, the CAR approach increased the uncertainty of the predictions (Fig. 4).

For central Finland, environmental variables in the non-spatial approach did not seem to make large differentiation between the known presence sites and sites without detection (Fig. 6) (which is not a modeling goal in itself, because a detection failure may arise either from to real absence, or from imperfect detection). The CAR models, however, were able to produce a more clear differentiation between the sites with presence and the sites without detection. This indicates that those sites represent real absences, which can be powerfully explained while accounting for the partial autocorrelation in the response. For Karelia, however, even the spatial approach did not differentiate the presence sites from the sites without detection (Fig. 6). The differentiation between the known presence sites and pseudo-absence sites was expectedly high for both areas.

Across all the models for both areas, residual autocorrelation remained significant up to a distance of 5 km. Residuals of the models showed highest autocorrelation, as expected, for the territory level non-spatial models (Moran's I values in the first distance class ranging between 0.33- 0.36 for central Finland, and 0.42 -0.67 for Karelia) (results not shown). The residual autocorrelation of the non-spatial multi-scale models was weaker than that of the territory level models (fist distance class autocorrelation ranging between 0.28 - 0.30 for central Finland, and 0.19 – 0.24 for Karelia). For central Finland, the CAR approach was able to reduce the spatial autocorrelation of the model residuals, compared to the non-spatial territory and multi-scale models (for example, Moran's I values for multi-scale CAR models in a first distance class ranging between 0.24-0.25). For Karelia, on the other hand, the switch from the non-spatial territory level models to multi-scale models resulted in a greater and more consistent reduction in residual autocorrelation than a switch from the non-spatial approach to the spatial one.

DISCUSSION

Our results suggests that the territory and landscape level habitat selection of the Siberian Jay shows differences in the two study areas, the distance between which is only 100 km in east-west direction (Tables 2 and 3). In central Finland, the Siberian Jay seemed to set stricter demands on its habitat, especially in its association to forest cover also at the landscape level. The occurrence of the species in central Finland was promoted by spruce mires, the occurrence of mire-forest mosaic (i.e., soft edges), young forests and increasing density of spruce-dominated forests at the territory level, and by forest cover and low human impact at the landscape level. In Karelia, the species occurrence increased with elevation, towards north, with spruce-dominated forests older than an average advanced thinning stand and with the density of spruce-dominated forests at the territory level, as well as with low human impact, young forests and mire areas at the landscape level.

At the territory level, variables associated with spruce cover (spruce mires in central Finland, and aged spruce-dominated forests in Karelia) were included in the best-screened models for Siberian jay occupancy. This is not surprising given the earlier research on the habitat associations (e.g., Väisänen et al. 1998, Edenius and Meyer 2002). However, our results suggest that presence of spruce dominated forests is not enough but the structural characteristics of spruce-dominated forests along with other territory and landscape level factors make a site suitable for the species.

In both areas, AIC_c based model selection gave support for the inclusion of the measure of the density of spruce-dominated forests (Appendices II and III). Forest thinning is being carried out 2-3 times during rotation period to remove competing, lower quality trees (Anonymous 2007). Intensive thinning results in rather uniform and open forest structure, compared with a forest stand managed with a lighter thinning. There are results showing that the female mortality rises with the age of the managed forest, because repeated thinning of the forest produces too open forest structure (Nystrand et al. 2010). Bergholm (2007) found out the total yearly harvest of pulpwood (instead of saw-timber) to be negatively associated with the Siberian Jay survival and immigration. The Siberian Jay seem to be a species that requires more dense forest structure than produced by the prevailing thinning recommendations (see e.g., Virkkala 1987, Griesser et al. 2007). Particularly, high density of spruce understorey (spruces < 15m) has been linked with high-quality jay territories (Sklepkovych 1997, Ekman et al. 1999, Ekman et al. 2001, Eggers et al. 2005, Angervuori 2008, Eggers et al. 2008). Unfortunately, the multi-source NFI-data does not allow account for the density of understorey separately from the prevailing (canopy) storey. This would, however, be important if guidelines for sustainable forestry practices within jay territories are to be formulated.

Somewhat surprising, the occurrence of the species was positively associated with the extent of young forests in both areas (also when the extent of forested area, irrespective of forest age, was controlled for) (Tables 2 and 3). This might be due to a positive association of the Siberian jay with young thinning stands of the age between 20-40 years, and not so much due to association with open clear-cuts and sapling stands younger than 20 year. The 20-40 years old young forests that have not yet undergone any thinning practices might offer dense forest structure, and also possibly provide nutrient-rich edge zones when associated with older spruce-dominated forests. But as the species may show a lagged response to environmental changes (discussed later), it is difficult to conclude whether the Siberian jay occurrence pattern showed an updated relationship with the full extent of young forests at the territory or landscape level. Whatever the true relationship of the Siberian Jay occupancy to young forests is, our results show that the recent territories are located in intensively managed forest areas and landscapes.

Our results provide only information on the territory and landscape level habitat selection, and not on the selection at the level finer than that. This should be kept in mind when the results are interpreted. The nest-site selection within territory is also an important level determining species occurrence patterns and reproductive output, but our data did not allow us to work on that level. The experimental feeding sites only indicate the locations of territories, but tell nothing about those of nesting sites.

Our results suggest that in central Finland the length of soft edges (mires/fens sharing border with forests) at the territory level was in positive association with the occurrence of the Siberian Jay. Also earlier research has suggested soft forest edges beneficial for the species (Sklepkovych 1997). In Karelia, the species presence was promoted in landscapes rich in mire areas, and the mire coverage was a better predictor of the species presence than the edge length. Thus, the measures to promote persistence of the Siberian jay populations at the southern edge of its distribution should not only concentrate on the characteristics of forest stands, but provide protection for the peatlands interspersed in the forested landscapes as well.

For both areas, it was clear that the spatial error term in territory level models encompasses also the spatial structure of the missing landscape level variables (Appendices II and III, Table 2). For this reason, the CAR approach for territory level models was superior to the nonspatial approach in terms of model fit and performance. When the landscape level variables were included into the models, the residual spatial autocorrelation turned to insignificant in multi-scale models of Karelia, and decreased for those of central Finland. For central Finland, the CAR approach was able to remarkably improve the model fit and performance of the multi-scale models (Appendix III and Table 3), although the spatial autocorrelation in model residuals still extended as significant up to 5 km. The further removal of the residual autocorrelation would possibly require a neighborhood size larger than the eight adjacent cells we have used. A larger neighborhood size would also possibly improve model performance and allow more safe inference, but it comes with the cost of an increasing computational intensity. Our results suggest that the distribution pattern of the Siberian Jay in central Finland cannot be fully understood and predicted on the basis of environmental covariates alone, but the social system of the species possibly leads to intrinsically autocorrelated distribution pattern. We cannot, however, rule out the possibility of missing covariates leading to extrinsic autocorrelative pattern, but based on the ecology of the species we consider the residual spatial autocorrelation more probably rising from intrinsic processes. The autocorrelated distribution pattern has also conservation implications: conservation measures for the species should also be aggregated in space to support the spatially organized social system of the species.

The probability of detecting the Siberian Jay in any given survey (given the species is present at a site) was, for both areas, *c.* 0.10 across all the models. This rather low detection probability may be also reflected to model performance (see MacKenzie et al. 2002, Reese et al. 2005), especially in Karelia when combined with a rather small sample size. We assumed that the detection probability for seeing at least one Siberian Jay individual in an occupied territory is constant and independent from the detections of previous surveys. It is likely, however, that detection probability varies in time, with survey effort and among surveyors but our data did not allow inspection of these factors. The low estimated detection probability also suggests that any further study on the Siberian jay distribution pattern should take imperfect detection into account in order not to underestimate the species distribution, and to come up with correct conclusions of its habitat associations.

We performed the model selection procedure for nonspatial models, and then fitted the spatial counterparts for the best-screened models. This was a decision dictated by the long computation times needed to run spatial models as well as a lack of recommended approaches for conducting model selection on Bayesian models. We acknowledge that the relative importance of variables may change when non-spatial and spatial models are compared, and model selection should be ideally conducted on spatial models (see e.g., (Lichstein et al. 2002, Tognelli and Kelt 2004, Hoeting et al. 2006). We still not believe that there is a danger of having excluded important variables from the final models. It is only that the final models may now be too complex, including also some variables that might drop if the model selection would be done on spatial models only (Hoeting et al. 2006, Diniz-Filho et al. 2008).

Models were fitted and their performance was measured exploiting the whole data set. Although this enabled us to use full content of our original datasets to have as precise parameter estimates as possible, this decision also comes with the risk of overfitting. Cross-validation procedures would have provided solutions to evaluate the predictive power of the models beyond the data used for the model fitting, but computational intensity of these procedures turned this too time-consuming to carry out. MPA as a measure of model performance tells thus more about the classification accuracy of a model, but not on its predictive power in a strict sense. But as the habitat relationship of the Siberian jay seemed to be area-specific (based on the results from our two study areas), it is difficult to tell what the potential predictive range of the models is outside the study areas.

In this study we used occurrence data to shed light on the habitat selection criteria of the jays. As Siberian Jays may abandon their territories as a response to adverse changes within their territories only after a time lag (i.e., Griesser et al. 2007), the distribution pattern may give somewhat biased information on the habitat requirements of the species. Nesting success data would be needed to distinguish viable territories (sources) from occupied but unproductive ones (sinks), and such data would give more detailed information on the habitat characteristics that determine the success of the Siberian Jay in managed forests. There is some indication that territories producing offspring differ from the ones with no offspring production in terms of their landscape level forest characteristics (Angervuori 2008).

ACKNOWLEDGEMENT

The study was funded by Kone Foundation (as a grant to AL). We are grateful to the Central Finland Bird Club, Lintuyhdistys Kuikka, North Karelian Ornithological Society, Suomenselän Lintutieteellinen yhdistys ry, and Metsähallitus for providing us with the occurrence data of the Siberian Jay. We thank the Finnish Forest Research Institute for allowing us to use the multi-source NFI-data in our research.

REFERENCES

Anonymoys 2007. Hyvän metsänhoidon suositukset. - Metsäkustannus.

Anonymoys 2009. Finnish statistical yearbook of forestry. - Metsäntutkimuslaitos.

Angervuori, A. 2008. Metsämaiseman muutoksen vaikutus kuukkelin esiintymiseen Etelä-Pohjanmaalla. - MSc Thesis, University of Helsinki.

Auvinen, P. 1997. Metsänmittaus. - Opetushallitus.

Bahn, V., O'Connor, R. J. and Krohn, W. B. 2006. Importance of spatial autocorrelation in modeling bird distributions at a continental scale. - Ecography 29: 835-844.

Bergholm, T. 2007. Kuukkelin (Perisoreus infaustus) populaatioekologia eteläisellä Pohjanmaalla. - MSc Thesis, University of Helsinki.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. - J. Roy. Statistical Society, Series B 36: 192-236.

Betts, M. G., Diamond, A. W., Forbes, G. J., Villard, M. and Gunn, J. S. 2006. The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. - Ecol. Model. 191: 197-224.

Braunisch, V., Bollmann, K., Graf, R. F. and Hirzel, A. H. 2008. Living on the edge—Modelling habitat suitability for species at the edge of their fundamental niche. - Ecol. Model. 214: 153-167.

Buler, J. J., Moore, F. R. and Woltmann, S. 2007. A multi-scale examination of stopover habitat use by birds. - Ecology 88: 1789-1802.

Burnham, K. P., Anderson, D. R. and Burnham, K. P. 2002. Model selection and multimodel inference : a practical information-theoretic approach. - Springer.

Carroll, C., Johnson, D. S., Dunk, J. R. and Zielinski, W. J. 2010. Hierarchical Bayesian spatial models for multispecies conservation planning and monitoring. - Conserv. Biol. 24: 1538-1548.

Chefaoui, R. M. and Lobo, J. M. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. - Ecol. Model. 210: 478-486.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V. and Wikle, C. K. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. - Ecol. Appl. 19: 553-570.

Cushman, S. A. and McGarigal, K. 2002. Hierarchical, Multi-scale decomposition of species-environment relationships. - Landscape Ecol. 17: 637-646.

Diniz-Filho, J. A. F., Rangel,Thiago Fernando L. V. B. and Bini, L. M. 2008. Model selection and information theory in geographical ecology. - Global. Ecol. Biogeogr. 17: 479-488.

Dormann, C. F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. - Global. Ecol. Biogeogr. 16: 129-138.

Edenius, L. and Elmberg, J. 1996. Landscape level effects of modern forestry on bird communities in North Swedish boreal forests. - Landscape Ecol. 11: 325-338.

Edenius, L. and Meyer, C. 2002. Activity budgets and microhabitat use in the Siberian Jay Perisoreus infaustas in managed and unmanaged forest. - Ornis Fenn. 79: 26-33.

Eggers, S., Griesser, M. and Ekman, J. 2008. Predator-induced reductions in nest visitation rates are modified by forest cover and food availability. - Behav. Ecol. 19: 1056-1062.

Eggers, S., Griesser, M., Andersson, T. and Ekman, J. 2005. Nest predation and habitat change interact to influence Siberian jay numbers. - Oikos 111: 150-158.

Ekman, J., Bylin, A. and Tegelstrom, H. 1999. Increased lifetime reproductive success for Siberian jay (Perisoreus infaustus) males with delayed dispersal. - Proc. R. Soc. B.-Biol. Sci. 266: 911-915.

Ekman, J., Eggers, S. and Griesser, M. 2002. Fighting to stay: the role of sibling rivalry for delayed dispersal. - Anim. Behav. 64: 453-459.

Ekman, J., Eggers, S., Griesser, M. and Tegelstrom, H. 2001. Queuing for preferred territories: delayed dispersal of Siberian jays. - J. Anim. Ecol. 70: 317-324.

Engler, R., Guisan, A. and Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. - J. Appl. Ecol. 41: 263-274.

Esseen, P., Ehnström, B., Ericsson, L. and Sjöberg, K. 1997. Boreal forests. - Ecol. Bull. 46: 16-47.

Farnsworth, M. L., Hoeting, J. A., Hobbs, N. T. and Miller, M. W. 2006. Linking chronic wasting disease to mule deer movement scales: a hierarchical Bayesian approach. - Ecol. Appl. 16: 1026-1036.

Fortin, M. and Dale, M. 2005. Spatial analysis. A guide for ecologists. - Cambridge University Press.

Gienapp, P. and Merilä, J. 2011. Sex-specific fitness consequences of dispersal in Siberian jays. - Behav. Ecol. Sociobiol. 65: 131-140.

Graf, R. F., Bollmann, K., Suter, W. and Bugmann, H. 2005. The importance of spatial scale in habitat models: capercaillie in the Swiss Alps. - Landscape Ecol. 20: 703-717.

GRASS Development Team. 2010. Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.0. Open Source Geospatial Foundation.

Griesser, M., Nystrand, M. and Ekman, J. 2006. Reduced mortality selects for family cohesion in a social species. - Proc. R. Soc. B-Biol. Sci. 273: 1881-1886.

Griesser, M., Nystrand, M., Eggers, S. and Ekman, J. 2007. Impact of forestry practices on fitness correlates and population productivity in an open-nesting bird species. - Conserv. Biol. 21: 767-774.

Gu, W. and Swihart, R. K. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. - Biol. Conserv. 116: 195-203.

Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. - Ecol. Model. 135: 147-186.

Heikkinen, R. 2000. Suojelualueverkon merkitys metsälajistolle : lehtojen putkilokasvit, metsien lahopuukovakuoriaiset, havu- ja sekametsien linnut. - Suomen ympäristökeskus.

Helle, P. and Jarvinen, O. 1986. Population trends of north Finnish land birds in relation to their habitat selection and changes in forest structure. - Oikos 46: 107-115.

Hines, J.E. 2001. PRESENCE2 - Software to estimate patch occupancy and related parameters. - USGS-PWRS. http://www.mbr-pwrc.usgs.gov/software/presence.html.

Hirzel, A. H., Hausser, J., Chessel, D. and Perrin, N. 2002. Ecological niche factor analysis: How to compute habitat-suitability maps without absence data? - Ecology 83: 2027-2036.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C. and Guisan, A. 2006. Evaluating the ability of habitat suitability models to predict species presences. - Ecol. Model. 199: 142-152.

Hoeting, J. A., Davis, R. A., Merton, A. A. and Thompson, S. E. 2006. Model Selection For Geostatistical Models. - Ecol. Appl. 16: 87-98.

Ishihama, F., Takeda, T., Oguma, H. and Takenaka, A. 2010. Comparison of effects of spatial autocorrelation on distribution predictions of four rare plant species in the Watarase wetland. - Ecological Research 25: 1057-1069.

Järvinen, O., Kuusela, K. and Väisänen, R. A. 1977. Effects of modern forestry on the numbers of breeding birds in Finland in 1945.1975. - Silva Fennica 11: 284-294. (In Finnish with English summary).

Kemppainen, J. and Kemppainen, O. 1991. Kuukkelin esiintymisestä Etelä-Suomessa 1960-1990. - Lintumies 1: 20-29.

Kouki, J., Väänänen, A. 2000. Impoverishment of resident old-growth forest bird assemblages along isolation gradient of protected areas in eastern Finland. - Ornis Fenn. 77: 145-154.

Latimer, A. M., Wu, S., Gelfand, A. E. and Silander Jr., J. A. 2006. Building statistical models to analyze species distributions. - Ecol. Appl. 16: 33-50.

Legendre, P. 1993. Spatial autocorrelation - trouble or new paradigm. - Ecology 74: 1659-1673.

Levin, S. A. 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. - Ecology 73: 1943-1967.

Lichstein, J. W., Simons, T. R., Shriner, S. A. and Franzreb, K. E. 2002. Spatial autocorrelation and autoregressive models in ecology. - Ecol. Monogr. 72: 445-463.

Lillandt, B. G. 1993. Lavskrikans (Perisoreus infaustus) populationsutveckling inom ett sammanhängade skogsområde i Sydösterbotten 1974-1992. - MSc Thesis, University of Helsinki.

Lillandt, B. G. 2000. Suupohjan kuukkelitutkimus 27 vuotta 1974-2000. - Hippiäinen 30: 13-26.

Link, W. A. and Barker, R. J. 2010. Bayesian Inference: with ecological applications. - Elsevier, Academic Press.

Louhia, O. 2001. Kuukkelin (*Perisoreus infaustus*) esiintyminen ja runsaus suhteessa metsämaiseman rakenteeseen. - MSc Thesis, University of Oulu.

Löfman, S. and Kouki, J. 2001. Fifty years of landscape transformation in managed forests of southern Finland. - Scand. J. For. Res. 16: 44-53.

MacKenzie, D. I. 2006. Occupancy estimation and modeling : inferring patterns and dynamics of species. - Elsevier Academic Press.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A. and Langtimm, C. A. 2002. Estimating site occupancy rates when detection probabilities are less than one. - Ecology 83: 2248-2255.

McGarigal, K., Cushman, S. A., Neel M. C. and Ene, E. 2002. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: www.umass.edu/landeco/research/fragstats/fragstats.html

Miller, J., Franklin, J. and Aspinall, R. 2007. Incorporating spatial dependence in predictive vegetation models. - Ecol. Model. 202: 225-242.

Muukkonen, P., Takala, T. and Virtanen, T. 2009. Differences in the forest landscape structure along the Finnish-Russian border in southern Karelia. - Scand. J. For. Res. 24: 140-148.

Naugle, D. E., Higgins, K. F., Estey, M. E., Johnson, R. R. and Nusser, S. M. 2000. Local and landscape-level factors influencing black tern habitat suitability. - J. Wildl. Manage. 64: 253-260.

Nystrand, M., Griesser, M., Eggers, S. and Ekman, J. 2010. Habitat-specific demography and source–sink dynamics in a population of Siberian jays. - J. Anim. Ecol. 79: 266-274.

Östlund, L., Zackrisson, O. and Axelsson, A. L. 1997. The history and transformation of a Scandinavian boreal forest landscape since the 19th century. - Can. J. Forest Res. 27: 1198-1206.

Pihlajaniemi, M. 2006. Kuukkeli Etelä-Suomessa. - Metsähallituksen luonnosuojelujulkaisuja. Sarja A 158.

Rassi, P., Hyvärinen, E., Juslén, A. and Mannerkoksi, I. (Eds.) 2010. The 2000 Red List of Finnish Species. Ympäristöministeriö, Suomen Ympäristökeskus, Helsinki (in Finnish with English summary)

Reese, G. C., Wilson, K. R., Hoeting, J. A. and Flather, C. H. 2005. Factors affecting species distribution predictions: a simulation modeling experiment. - Ecol. Appl. 15: 554-564.

Royle, J. A. and Dorazio, R. M. 2008. Hierarchical modeling and inference in ecology : the analysis of data from populations, metapopulations and communities. - Academic Press.

Royle, J. A., Kery, M., Gautier, R. and Schmid, H. 2007. Hierarchical spatial models of abundance and occurrence from imperfect survey data. - Ecol. Monogr. 77: 465-481.

Sklepkovych, B. 1997. Kinship and conflict - Resource competition in a proto-cooperative species: the Siberian Jay. - Dissertation, University of Stockholm.

Teeffelen, A. J. A., and Ovaskainen, O. 2007. Can the cause of aggregation be inferred from species distributions? - Oikos 116: 4-16.

Thomas, A., O'Hara, B., Ligges, U, and Sturtz, U. 2006. Making BUGS Open. - R News 6: 12-17.

Tikkanen, O., Heinonen, T., Kouki, J. and Matero, J. 2007. Habitat suitability models of saproxylic red-listed boreal forest species in long-term matrix management: Cost-effective measures for multi-species conservation. - Biol. Conserv. 140: 359-372.

Tognelli, M. F. and Kelt, D. A. 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. - Ecography 27: 427-436.

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K. and Possingham, H. P. 2003. Improving precision and reducing bias in biological surveys: estimating false negative error rates. - Ecol. Appl. 13: 1790-1801.

Uimaniemi, L., Orell, M., Mönkkönen, M., Huhta, E., Jokimäki, J. and Lumme, J. 2000. Genetic diversity in the Siberian jay *Perisoreus infaustus* in fragmented old-growth forests of Fennoscandia. - Ecography 23: 669-677.

Väisänen, R. A., Lammi, E., Koskimies, P. and Kostet, J. 1998. Muuttuva pesimälinnusto. - Otava.

Virkkala, R. 1987. Effects of forest management on birds breeding in northern Finland. - Ann. Zool. Fenn. 24: 281-294.

Wiens, J. A. 1989. Spatial Scaling in Ecology. - Funct. Ecol. 3: 385-397.

Wikle, C. K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. - Ecology 84: 1382-1394.

Wilson, T. L., Odei, J. B., Hooten, M. B. and Edwards Jr, T. C. 2010. Hierarchical spatial models for predicting pygmy rabbit distribution and relative abundance. - J. Appl. Ecol. 47: 401-409.

Zaniewski, A. E., Lehmann, A. and Overton, J. M. C. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. - Ecol. Model. 157: 261-280.

TABLE 1 The territory and landscape level variables used in the modelling.

Variable	Abbreviation	Unit
Territory level, r = 750m		
Average elevation	Elev	meters (above sea level)
Edge length shared between forests (>100m^3/ha) and mires/fens	Edge	m/ha
Forests >100m^3/ha	Forest100_T	% of total area
Forests < 40 years	Young_T	% of total area
Human-dominated environment	Human_T	% of total area
Maximum density of spruce-dominated forests	Densitymax	the ratio of basal area to average stock height
Median density of spruce-dominated forests	Densitymed	the ratio of basal area to average stock height
Mires/fens	Mire_T	% of total area
Spruce mire forests >100m^3/ha	Smire_T	% of total area
Sprude-dominated forests > 100m^3/ha	Spruce100_T	% of total area
Spruce-dominated forests > 170m^3/ha	Spruce170_T	% of total area
Topographic variation of the terrain	Topo	SD of the elevation
X-coordinate	Xcoord	10 km (from the study area border)
Y-coordinate	Ycoord	10 km (from the study area border)
Landscape level, r = 5km		
Forests < 40 years	Young_L	% of total area
Forests >100m3/ha	Forest100_L	% of total area
Human-dominated environment	Human_L	% of total area
Mires/fens	Mire_L	% of total area
Spruce-dominated forests > 100m^3/ha	Spruce100_L	% of total area
Spruce-dominated forests > 170m^3/ha	Spruce170_L	% of total area
Area-weighted mean patch size of the spruce-dominated forests > 170m3/ha	Spruce170_awm	ha

TABLE 2 The Bayesian CAR-model parameter estimates (the posterior sample mean and standard deviation) for the best-screened territory-level models. If the 95 % confidence interval does not encompass zero, the parameters is marked with asterisk (otherwise the Bayesian Confidence Interval, BCI, is reported). Detection probability gives a parameter estimate for the probability of observing a Siberian jay during a survey given a site was occupied. Minimal Predicted Area stands for the areal proportion (%) of the study area when the predicted probability of species occurrence was set to exclude 10% of the most the improbable occurrences.

	central Finland				Karelia		
Variable	T1	T2	T3	T4	T1	T2	T3
Elev					0.007 (0.002)*	0.007 (0.002)*	
Edge, linear term	0.05 (0.02)*	0.06 (0.02)*					
Young_T	0.11 (0.03)*	0.13 (0.03)*	0.08 (0.03)*	0.09 (0.03)*		0.10 (0.06)2)	0.12 (0.04)*
Human_T			-0.06 (0.03)*	-0.06 (0.03)*			
Densitymax						5.98 (3.23)3)	4.40 (2.44)6)
Densitymed		6.62 (3.11)*		5.30 (2.85)1)			
Mire_T, linear term						0.15 (0.07)*	0.19 (0.05)*
Mire_T,quadratic term						-0.004 (0.003)4)	-0.005 (0.002)*
Smire_T	0.62 (0.18)*	0.52 (0.18)*	0.67 (0.16)*	0.63 (0.17)*			
Spruce170_T						0.18 (0.10)5)	0.19 (0.08)*
Ycoord					0.34 (0.15)*	0.48 (0.21)*	0.37 (0.14)*
Detection probability	0.12	0.12	0.12	0.12	0.10	0.10	0.12
Deviance (sample mean)	871.1	865.1	865.6	870.0	239.9	231.8	251.9
Minimum Predicted Area	19.8	24.7	21.2	24.0	44.3	42.0	41.0

1) BCI [-0.067; 11.06]
2) BCI [-0.02; 0.23]
3) BCI [-0.28; 12.83]
4) BCI [-0.010; 0.0003]
5) BCI [-0.009; 0.382]
6) BCI [-0.32; 9.30]

TABLE 3 The Bayesian CAR-model parameter estimates (the posterior sample mean and standard deviation) for the best-screened multi-scale models. If the 95 % confidence interval does not encompass zero, the parameters is marked with asterisk (otherwise the Bayesian Confidence Interval, BCI, is reported). Minimal Predicted Area stands for the areal proportion (%) of the study area when the predicted probability of species occurrence is s to exclude the 10% of the most improbable occurrences.

	central Finland				Karelia		
Variable	L1	L2	L3	L4	L1	L2	L3
Territory level							
Elev					0.006 (0.002)*	0.007 (0.003)*	0.006 (0.002)*
Edge, linear term	0.04 (0.02)$^{1)}$	0.04 (0.03)$^{2)}$					
Young_T	0.11 (0.03)*	0.12 (0.03)*	0.09 (0.03)*	0.10 (0.03)*			
Densitymax		5.4 (2.8)$^{3)}$		4.14 (2.81)$^{5)}$	7.84 (4.03)*	8.34 (4.51)*	
Smire_T	0.52 (0.17)*	0.54 (0.19)*	0.62 (0.17)*	0.62 (0.18)*			
Spruce170_T					0.20 (0.10)*	0.18 (0.11)$^{8)}$	0.12 (0.08)$^{10)}$
Ycoord					0.40 (0.21)*	0.39 (0.27)$^{9)}$	0.35 (0.17)*
Landscape level							
Young_L					0.32 (0.10)*	0.35 (0.12)*	0.27 (0.09)*
Forest100_L	0.08 (0.04)*	0.09 (0.04)*	0.08 (0.04)*	0.09 (0.04)*			
Human_L		-0.06 (0.05)$^{4)}$	-0.10 (0.05)*	-0.09 (0.05)$^{6)}$	-0.40 (0.17)*	-0.46 (0.20)*	-0.24 (0.12)*
Mire_L, linear term					0.22 (0.09)*	0.24 (0.09)*	0.23 (0.07)*
Mire_L, quadratic term					-0.002 (0.004)$^{7)}$		
Detection probability	0.12	0.12	0.12	0.12	0.10	0.10	0.10
Deviance (sample mean)	867.0	871.7	868.7	872.8	227.5	225.3	227.0
Minimum Predicted Area	19.7	21.3	19.7	20.1	41.9	38.7	43.9

1) BCI [-0.002; 0.088]
2) BCI [-0.01; 0.09]
3) BCI [-0.06; 11.0]
4) BCI [-0.17; 0.04]
5) BCI [-1.30; 9.61]
6) BCI [-0.184; 0.002]
7) BCI [-0.011; 0.005]
8) BCI [-0.01; 0.40]
9) BCI [-0.12; 0.96]
10) BCI [-0.03; 0.12]

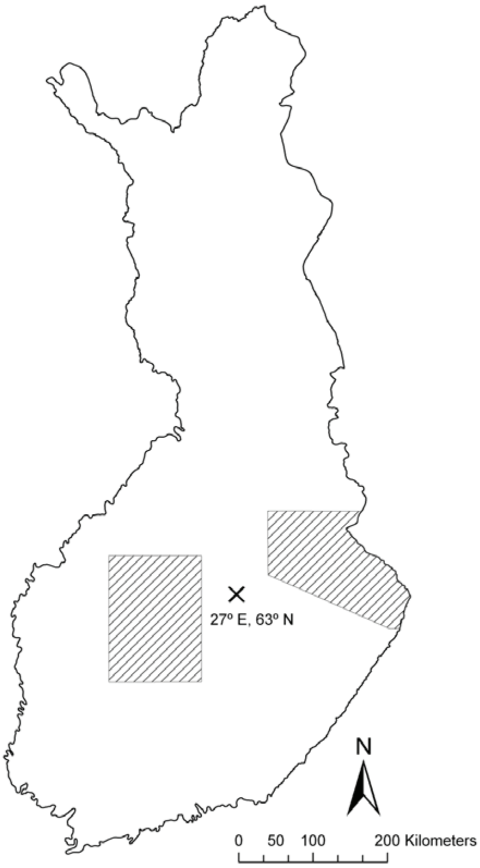


FIGURE 1 Our two study areas, one located in central Finland and the other in Karelia.

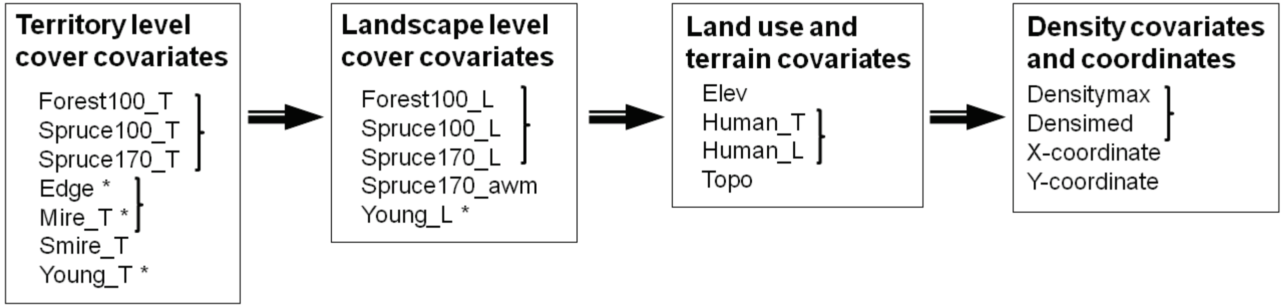


FIGURE 2 The chart for the order in which the variables were allowed to enter the models. Brackets indicate a group of variables that were not allowed to occur simultaneously in the models. Asterisks indicate the variables for which we also tested the quadratic relationships. For the territory-level models, the landscape level cover covariates were omitted.

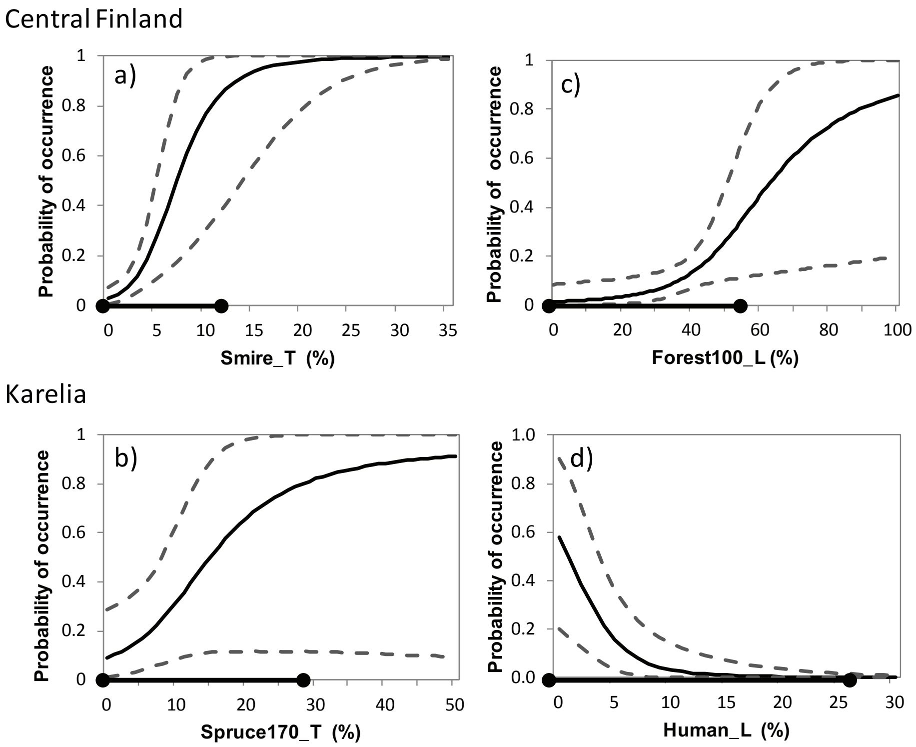


FIGURE 3 Probability of the Siberian Jay occurrence in central Finland, based on the CAR L1 model, as a function of the covariates Smire_T (a) and Forest100_L (c), as well as in Karelia, based on the CAR L2 model, as a function of the covariates Spruce170_T (b) and Human_L (d). The other environmental covariates in the models were fixed to their mean values. The solid line represents the mean of the posterior predictions and the broken lines represent the 95 % confidence interval for the estimate. The observed range of variation for the sites with the occurrence data is represented with a strong solid line bounded by the circles.

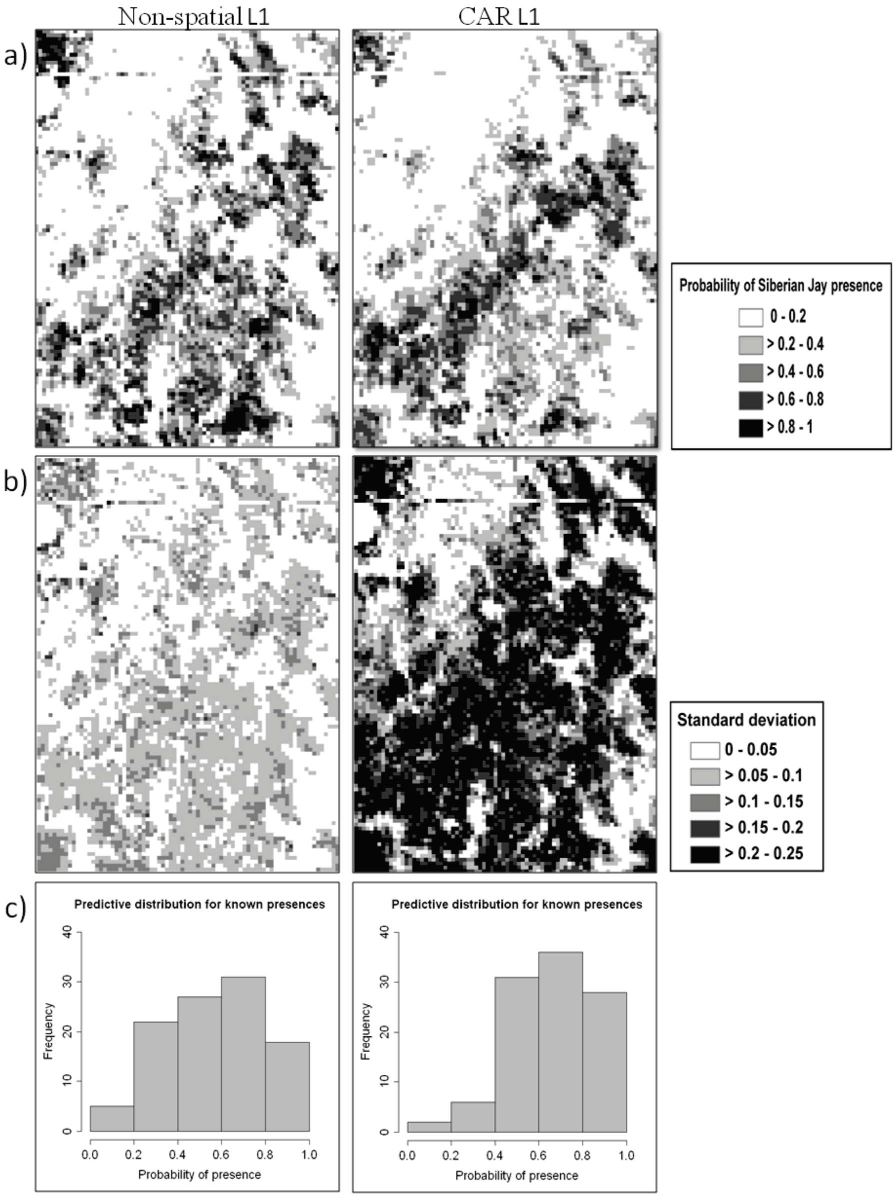


FIGURE 4 Moran's I values of the first eight distance classes (bin width of 2.5 km, ranging up to 20 km) for the three modeling methods; nonspatial (black line), CAR (grey line) and autologistic regression (dashed line).

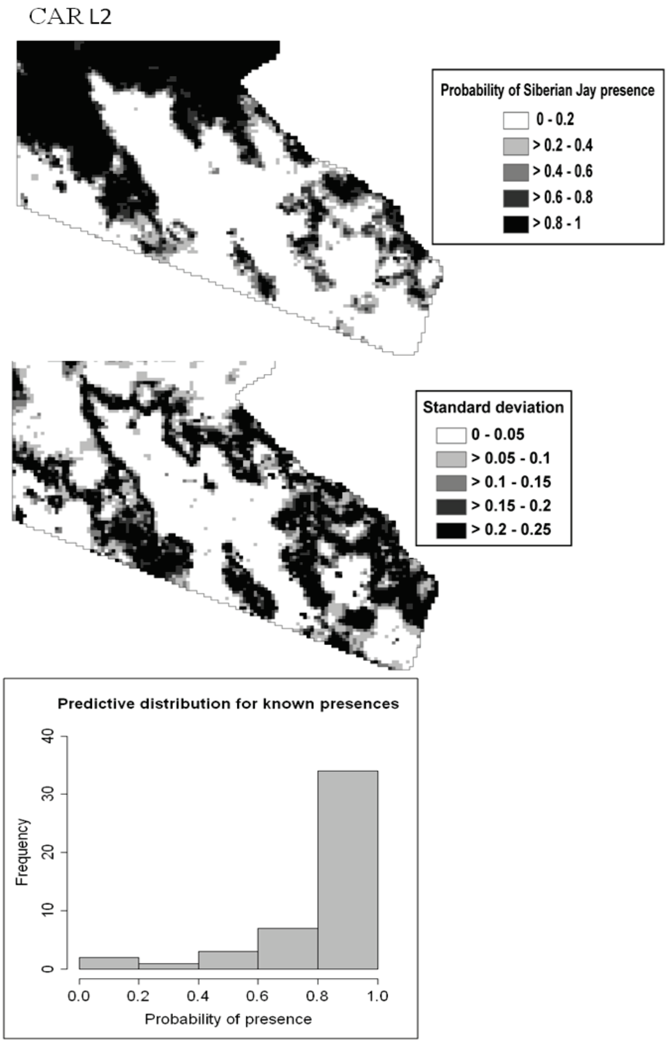


FIGURE 5 Predicted probabilities for the presence of the Siberian Jay (upper-most panel) in Karelia based on the multi-scale, spatial CAR L2 model. For each grid cell, predicted probability represents the mean of the posterior sample. The middle-panel shows the standard deviations of the samples. The lower-most panel shows the distribution of predicted probabilities for the known presence sites (*n*=47) involved in the modelling.

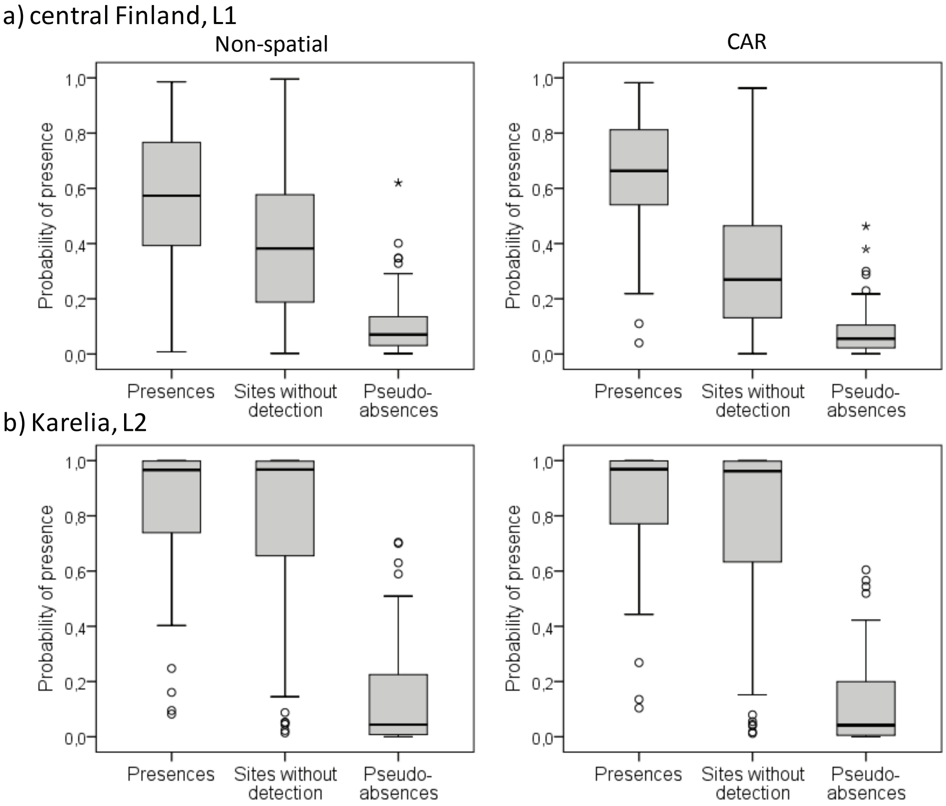


FIGURE 6 Distributions of predictive probabilities for the presence sites, the sites without observations (either non-occupied sites or sites with failed detection) and the pseudo-absence sites based on a non-spatial and spatial CAR model for central Finland (multi-scale L1 model) and Karelia (multi-scale L2 model).

Appendix I ENFA results for the two study areas, central Finland (a) and Karelia (b). Variance explained (%) by the first four factors, and coefficient values for the variables (sorted by their decreasing absolute values). The greater the absolute value of a variable on the marginality factor is, the more the habitat selection of the species deviates from the landscape mean (the sign indicating the direction of departure). The greater the absolute value of a variable on the specialization factor is, the narrower the species niche is compared to the global distribution (the sign is now arbitrary).

a) ENFA-results for the study area of central Finland

Marginality factor (18%)	Specialization factor		
	1 (36 %)	2 (14 %)	3 (12 %)
Smire (0.50)	Densitymax (0.89)	Human_T (0.64)	Spruce170_T (0.74)
Cover_T (0.46)	Spruce170_T (0.25)	Edge (0.41)	Elev (-0.43)
Human_T (-0.39)	Edge (0.25)	Cover_T (0.40)	Cover_T (-0.42)
Young_T (0.34)	Cover_T (-0.19)	Spruce170_T (-0.39)	Edge (0.26)
Spruce170_T (0.34)	Elev (-0.15)	Young_T (0.24)	Densitymax (-0.15)
Edge (0.33)	Smire (-0.11)	Elev (-0.15)	Smire (-0.05)
Elev (0.21)	Young_T (0.08)	Densitymax (0.14)	Human_T (0.05)
Densitymax (-0.02)	Human_T (-0.01)	Smire (0.03)	Young_T (-0.04)

b) ENFA-results for the study area of Karelia

Marginality factor (26%)	Specialization factor		
	1 (25 %)	2 (21 %)	3 (8%)
Elev (0.64)	Human_T (0.53)	Spruce170_T (-0.70)	Cover_T (-0.67)
Human_T (-0.60)	Edge (0.42)	Smire (0.57)	Smire (0.57)
Young_T (0.36)	Cover_T (0.40)	Human_T (-0.35)	Spruce170_T (0.34)
Spruce170_T (0.28)	Young_T (0.38)	Young_T (-0.15)	Elev (-0.27)
Edge (0.15)	Elev (0.32)	Edge (-0.12)	Densitymax (-0.15)
Densitymax (0.05)	Spruce170_T (-0.29)	Densitymax (0.11)	Human_T (-0.11)
Smire (0.01)	Smire (0.22)	Elev (0.08)	Young_T (0.04)
Cover_T (0.00)	Densitymax (-0.03)	Cover_T (0.06)	Edge (-0.03)

Appendix II The Bayesian non-spatial logistic model parameter estimates (the posterior sample mean and standard deviation) for the best-screened territory-level models. If the 95 % confidence interval does not encompass zero, the parameters is marked with asterisk (otherwise the BCI is reported). Minimal Predicted Area stands for the areal proportion (%) of the study area when the predicted probability of species occurrence thresholded to exclude the 10% of the most improbable occurrences.

	central Finland				Karelia		
Variable	T1	T2	T3	T4	T1	T2	T3
Elev					0.007 (0.001)*	0.006 (0.002)*	
Edge, linear term	0.04 (0.02)*	0.05 (0.02)*					
Young_T	0.11 (0.03)*	0.12 (0.03)*	0.07 (0.02)*	0.08 (0.02)*		0.09 (0.05)2)	0.11 (0.04)*
Human_T			-0.05 (0.02)*	-0.06 (0.02)*			
Densitymax						5.21 (2.78)*	4.46 (2.19)*
Densitymed		5.16 (2.17)*		3.70 (2.15)1)			
Mire_T, linear term						0.13 (0.06)*	0.17 (0.05)*
Mire_T,quadratic term						-0.004 (0.002)3)	-0.005 (0.002)*
Smire_T	0.61 (0.14)*	0.56 (0.15)*	0.66 (0.14)*	0.65 (0.15)*			
Spruce170_T						0.15 (0.09)4)	0.17 (0.07)*
Ycoord					0.34 (0.12)*	0.45 (0.15)*	0.28 (0.10)*
Detection probability	0.11	0.11	0.11	0.11	0.10	0.10	0.12
Deviance (sample mean)	952.2	947.5	951.5	949.4	243.0	240.2	252.2
DIC	957.2	953.3	956.5	955.3	246.9	248.8	260.3
AIC_c	958.5	955.0	958.3	957.3	249.2	251.2	262.97
Minimum Predicted Area	43.0	40.5	47.9	47.1	50.1	47.9	46.5

1) BCI [-0.28; 8.03]

2) BCI [-0.006; 0.197]

3) BCI [-0.008; 0.0004]

4) BCI [-0.006; 0.344]

Appendix III The Bayesian nonspatial logistic model parameter estimates (the posterior sample mean and standard deviation) for the best-screened multi-scale models. If the 95 % confidence interval does not encompass zero, the parameters is marked with asterisk (otherwise the BCI is reported). Minimal Predicted Area stands for the areal proportion (%) of the study area when the predicted probability of species occurrence is thresholded to exclude the 10% of the most improbable occurrences.

	central Finland				Karelia		
Variable	L1	L2	L3	L4	L1	L2	L3
Territory level							
Elev					0.006 (0.002)*	0.006 (0.002)*	0.005 (0.002)*
Edge, linear term	0.04 (0.02)*	0.03 (0.02)$^{1)}$					
Young_T	0.11 (0.02)*	0.11 (0.03)*	0.08 (0.02)*	0.09 (0.02)*			
Human_T							
Densitymax		3.80 (2.21)$^{2)}$		3.19 (2.20)$^{4)}$	7.57 (4.06)*	7.11 (3.83)*	
Smire_T	0.54 (0.15)*	0.57 (0.17)*	0.65 (0.16)*	0.67 (0.17)*			
Spruce170_T					0.20 (0.10)*	0.17 (0.09)*	0.12 (0.08)$^{6)}$
Ycoord					0.41 (0.21)*	0.41 (0.19)*	0.37 (0.18)*
Landscape level							
Young_L					0.31 (0.11)*	0.30 (0.10)*	0.26 (0.09)*
Forest100_L	0.07 (0.02)*	0.08 (0.03)*	0.06 (0.03)*	0.07 (0.03)*			
Human_L		-0.06 (0.04)$^{3)}$	-0.09 (0.04)*	-0.09 (0.04)*	-0.39 (0.17)*	-0.38 (0.16)*	-0.23 (0.12)*
Mire_L, linear term					0.22 (0.09)*	0.20 (0.07)*	0.22 (0.07)*
Mire_L, quadratic term					-0.002 (0.004)$^{5)}$		
Detection Probability	0.10	0.10	0.11	0.11	0.10	0.10	0.10
Deviance (sample mean)	945.3	940.9	943.4	942.6	229.4	228.1	230.0
DIC	951.2	948.9	949.4	949.7	239.5	236.8	237.6
AIC$_c$	953.1	950.5	951.0	951.0	241.4	239.5	240.2
Minimum Predicted Area	36.9	39.2	39.2	39.8	41.7	40.8	43.8

1) BCI [-0.006; 0.075]
2) BCI [-0.57; 8.17]
3) BCI [-0.143; 0.015]
4) BCI [-1.19; 7.57]
5) BCI [-0.010; 0.006]
6) BCI [-0.02; 0.28]

IV

ACCOUNTING FOR SPACE IN SPECIES DISTRIBUTION MODELING – ALTERNATIVE ROUTES LEADING TO A SAME DESTINATION?

by

Anne Laita & Mikko Mönkkönen

Manuscript

ACCOUNTING FOR SPACE IN SPECIES DISTRIBUTION MODELING – ALTERNATIVE ROUTES LEADING TO A SAME DESTINATION?

Anne Laita & Mikko Mönkkönen

Department of Biological and Environmental Science, PO Box 35 (Survontie 9), 40014
University of Jyväskylä, Finland

ABSTRACT

The importance of spatially explicit approach in species distribution modeling is undisputed, because it allows for safe inference and enhanced predictive outcome. However, spatial approaches show also deficiently understood differences in their parameter estimations and modeling outcomes. The aim of this study was two-fold. First, we evaluated the inference and performance of two spatially explicit Bayesian modeling approaches, autologistic model and conditional autoregressive model (CAR), in order to see whether they end up with consistent inferences and predictive outcome. Second, we wanted to examine whether the broad-scale representation of space, i.e. landscape connectivity, could provide either an alternative spatial approach or additional explanatory power to spatial modeling. We used data on Siberian Jay (*Perisoreus infaustus*) occurrence in boreal forest landscapes in central Finland. According to our results, CAR and autologistic modeling approaches led to seemingly inconsistent inferences due to their different allocations between environment and space. They still both produced enhanced predictive performance compared to the non-spatial approach, albeit with somewhat deviating spatially projected predictions. Landscape connectivity seemed to be able to improve the model fit and performance of the non-spatial models, but still lacked behind the spatially explicit methods. The ability of landscape connectivity to improve the performance of the non-spatial models gives, however, an indication of the intrinsic autocorrelation underlying the species occupancy pattern. Our results show the strong intertwining between environmental factors and space, which serves as a key to understand the properties and thereby rising outcomes of different modeling methods.

INTRODUCTION

Species distribution modeling aims at explaining the species occupancy state based on environmental factors that are either causally or indirectly involved in controlling the distribution (e.g., Guisan and Zimmermann, 2000; Guisan et al., 2006). Species distribution models (SDMs) have gained a strong methodological foothold in many theoretical and applied ecological questions, ranging from reserve design (e.g., Araújo and Williams, 2000) to the range predictions for invasive species (e.g., Dullinger et al., 2009).

Species distributions usually exhibit spatial autocorrelation so that nearby locations resemble each other more than those further apart. The most part of this similarity is explained by the spatial structure in the environmental covariates that are controlling the occurrence patterns of a species. This part of autocorrelation can be simply "governed" in distribution models by identifying the correct environmental covariates behind species distributions, and entering them to models as explanators. The remaining part of the spatial autocorrelation in response values, not accounted for by environmental explanators (and included in model residuals), may be an outcome of several factors. In addition to missing or incorrectly specified environmental factors, biological processes may result in aggregated occurrence patterns (over the control of environmental covariates) (for a list of both exogenous and biological, i.e., intrinsic, aggregation factors, see Miller et al., 2007). These aggregating biological processes include, for example, social behaviors such as conspecific attraction and limited dispersal.

If the model residuals violate the key assumption that the residuals are independent and identically distributed, the inferences drawn from the model may be invalid. Violation of model assumptions may lead to inflated Type I error rate via the underestimation of the standard error in the parameter estimates, and additionally, biased parameter estimates (Cressie, 1993; Legendre, 1993; Segurado et al, 2006; Dormann et al., 2007; Kühn, 2007; Bini et al., 2009; Hoeting, 2009). In such cases, spatially explicit modeling approaches, which account for this residual non-independence, provide a viable solution.

Species distribution models are often based on binary presence-absence data. Modeling results for such data have been found to be especially sensitive to the choice of the spatial modeling method also (Dormann et al., 2007). Quite subtle differences in spatial autocovariance approaches may have major implications for the results, although the approaches may, at first sight, seem to have rather redundant focus (Betts et al., 2009). Applications of several methods to a single dataset has demonstrated that regression coefficients from different methods may differ in terms of their magnitude and level of significance (e.g, Dormann et al., 2007; Kissling and Carl, 2008; Bini et al., 2009; Beale et al., 2010).Therefore, for robust inferences and predictions from species distribution modeling there is a need to evaluate and compare alternative modeling methods.

Different spatial modeling methods have been usually compared in terms of parameter estimation, as if they are able to end up with the true, known parameter estimates (e.g., Carl and Kühn, 2007; Dormann et al., 2007; Dormann, 2007; Beale et al., 2010). Species distribution models, however, have also aims beyond parameter estimation and inference because in many cases they are used to predict species distributions. The prediction of species distributions based on the combined effect of all model constituting terms and covariates forms another important modeling dimension alongside covariate-wise parameter estimation and inference. It would be important to evaluate and compare the different modeling methods with respect to these both modeling dimensions, because the benefits associated with a spatial modeling method may also vary depending on the modeling target.

The spatial modeling methods account for the correlation structure in the response using the neighborhood of influence and different kind of weighting schemes. This neighborhood, over which the neighboring sites are assumed to influence the site under focus, may be of various sizes. The influence of neighboring sites on the focal site can also be weighted, for example, by the distance separating the sites, or they can be assumed to exert equal influence on a focal site within a specified neighborhood. Most often, the size of a neighborhood and weighting scheme are probed in an explanatory manner so that the selection of spatial extent and weights minimizes the residual autocorrelation, or maximizes model performance (e.g., Bahn et al, 2006, Carroll et al., 2010). The appropriate neighborhood size may also be informed based on the spatial correlogram of model residuals, see e.g. Wintle and Bardos (2006). This exploratory approach, however, lacks the hypothesis about the process creating the pattern. If the correct process were identified and more directly modeled, it would possibly lead to more effective reduction in the residual autocorrelation and to enhanced predictive power.

Landscape connectivity has been shown (theoretically and empirically) to be positively associated with the probability of patch occupancy (e.g., Bauerfeind et al., 2009; Brouwers and Newton, 2009; Ranius et al., 2010). Landscape connectivity may be viewed as a potential for the intrinsic biological processes (mediated by dispersal) to operate. The connectedness of a patch to other habitat patches allows the supply of individuals from the occupied patches within a species dispersal range, leading to enhanced colonization rate and depressed extinction rate (via rescue effect, Brown and Kodric-Brown, 1977). Intrinsic aggregation (and subsequent residual autocorrelation) may result from spatially-structured dispersal between areas, so that source populations supply individuals to their surroundings. The supply increases with the increasing size of a source patch and decreasing distance to the receiving area. This broad-scale landscape connectivity is seldom used in species distribution modeling, because the spatial effects are usually implicitly addressed via the neighborhood weighting of spatial methods. Landscape connectivity could, however, have a potential to explain and predict species distributions, and it might serve as a potential candidate for capturing intrinsic spatial effects for a species limited by dispersal in fragmented landscapes.

Here, we first compare the modeling outcome for two different spatially explicit modeling approaches with Bayesian implementation: autologistic regression and conditional autoregressive model (CAR). We contrast their parameter estimates and modeling outcomes with the results from the Bayesian non-spatial logistic regression model. We compare the modeling approaches both in terms of their explanatory and predictive performance. To our knowledge, this is the first time the modeling outcomes of autologistic regression and CAR modeling with Bayesian implementation are being compared, as previous comparisons have focused on the frequentist implementations of the methods (see e.g., Dormann, et al., 2007). Second, we evaluate whether the more process-based modeling of the aggregation would provide additional explanatory power to the traditionally used spatial methods with a neighborhood weighting, or alternatively, a new, competitive stand-alone way of removing autocorrelation and improving model fit. This is accomplished by directly including landscape connectivity as a variable into the models. Our modeling is based on the Siberian Jay (*Perisoreus infaustus*) occurrence data and data on forest characteristics at the territory and landscape level originating from a boreal forest landscape in Central Finland. Our study provides an example of a typical modeling case; complicatedly controlled species occurrence with no reference for the true parameter values and no full-coverage information for its occupancy. From practical perspective, the aim is to improve our ability to correctly model the distribution of this regionally endangered species.

METHODS

Autologistic regression model

Autologistic regression model aims at estimating how the binomial response value of a site (i.e. species occupancy) is affected by the response values of surrounding sites (i.e. occupancy status of the neighbors). The response values of the neighbors are added as an additional explanatory variable to a model, and the relationship of this variable to the response variable is specified by the associated extra parameter, the autologistic coefficient (Augustin et al. 1996; Dormann et al., 2007). Autologistic regression model has the formula:

$$logit(psi_i) = \alpha + \sum_k \beta_k x_{ki} + \beta_{k+1} \frac{\sum_{j \neq i} w_j z_j}{\sum w_j} + \varepsilon_i,$$

where α is the intercept; β_k is the coefficient for the explanatory variable x_k; $\beta_{k+1}$is the autologistic coefficient; z_j is the response values of a site *j*; w_j is the weight give to a site *j*'s influence over site *i*; and ε_i is the random effect term for a site *i*. The weights can be either binary weights (0/1) defining the neighborhood for a site *i* over which the neighbors influence its response value or follow some other weighting scheme, e.g. the influence of site j over i can be weighted as a function of geographical distance between the sites.

It is straightforward to estimate the autologistic term when there exists rather complete knowledge for species occurrence (i.e. occupancy status is known for sites under focus and their neighbors as defined by the neighborhood scheme), for example in a case of full-coverage atlas data. However, when the dataset contains many missing values, the estimation of the autologistic term turns out to be more complicated. In data-limited cases the autologistic term cannot be accessed directly, but can be estimated via an iterative procedure (as introduced by Augustin et al., 1996, see also Augustin et al., 1998). In this case, the autocovariate term is estimated based on the predicted response values (or alternatively, the values for the predicted probability of presence) rather than observed ones. This iterative procedure has been applied in several studies, for example in Osborne et al., 2001 and Piorecky and Prescott, 2006. The power of Bayesian analysis is highlighted in this kind of knowledge-limited case because the occupancy status of the missing sites is treated as underlying latent state variable, which is sampled from posterior distribution in a same manner as any other model parameter (see e.g., Wintle and Bardos, 2006; Royle and Dorazio, 2008).

The autologistic regression model has been shown to lead to enhanced predictive performance, compared to non-spatial models (e.g., Augustin et al., 1996; Hoeting et al., 2000; Osborne et al., 2001; Wintle and Bardos, 2006). There exists, however, rather contradicting evidence of the ability of autologistic regression models to correctly estimate model parameters. Some studies have reported autologistic regression models consistently underestimating the effects of environmental covariates, leading to even more biased parameter estimates than the ones produced by non-spatial models. The most convincing evidence for this underestimation comes from the simulation studies, where the true parameter value for a covariate is known (e.g., Dormann et al., 2007; Dormann, 2007a). Especially, a strong spatial structure in a covariate is found to be linked with biased parameter estimates (Betts et al., 2009). There exists, however, also support for autologistic regression models showing that non-biased parameter estimates have been reached with this method (e.g., Wu and Huffer, 1997; Huffer and Wu, 1998).

Conditional autoregressive (CAR) models

Conditional autoregressive (CAR) model incorporates autocorrelation in a spatially structured error term. In CAR model, probability of species presence depends, along with effect of environmental covariates, on a site-specific spatial random effect. The spatially structured variation in response is specified via an intrinsic Gaussian conditional autoregressive model (Besag, 1974). CAR model has the following formula:

$$logit(psi_i) = \alpha + \sum_k \beta_k \, x_{ki} + \rho_i + \varepsilon_i,$$

where ρ_i is the spatial random effect term for the site i (see the autologistic model formula above for the explanation of other model terms). The spatial random effect term of a site i (ρ_i) is dependent on the ρ terms of the neighboring sites. The neighborhood for a site i is specified by a researcher. The influence of neighboring sites on the spatial random effect of site i can also be weighted by their distance, but usually the neighboring sites are assumed to exert equal influence (as in the definition below). The following conditional relationship is assumed for the distribution of ρ_i:

$$\rho_i \mid \rho_{i+} \approx \text{Normal}\left(\frac{\sum_{j \in \delta i} \rho_j}{n_{i+}}, \frac{\sigma_\rho^2}{n_{i+}}\right),$$

where $i+$ stands for a set of neighbors of site i;site j belongs to the neighborhood of site i (denoted by δi); n_{i+} denotes the number of sites which are neighbors of site i; and σ_ρ^2 is the variance for all sites. Thus, the random effect for site i follows a normal distribution with mean corresponding the average random effect value of its neighbors, and with variance inversely proportional to the number of neighbors.

Bayesian CAR models in the modeling of binary occurrence data have been applied in a few, quite recent studies (e.g., Farnsworth et al., 2006; Latimer et al., 2006; Carroll et al., 2010; Ishihama et al., 2010). Beale et al. (2010) demonstrated that a Bayesian implementation of the conditional auto-regressive model (on data with normally distributed errors) performed well when it was evaluated against simulated data with known properties. Like autologistic regression, CAR modeling has been shown to possess superior predictive capability over the non-spatial modeling (e.g., Carroll et al., 2010).

Comparisons of the modeling methods

Our datasets and model selection procedure are described in detail in Laita and Mönkkönen (unpublished). In this earlier study we showed that CAR modeling was superior to non-spatial counterpart in terms of model fit and predictive performance. The spatial models including also landscape level variables produced better results than the models including only territory level variables, giving support to the importance of landscape context in shaping the species distribution pattern. The model we are relying in this study is the best-performing model from our previous study. The model includes three territory-level variables (T) and one landscape level variable (L). The territory level variables provide information of the characteristics of forests at the territory scale (*ca*. 2 km^2), while the landscape level variable describes the overall availability of forests in the landscape within the average 5 km dispersal distance of the species (ca. 80 km^2, see Gienapp and Merilä, 2011). In the modeling, we also controlled for the imperfect detection that is often inherent in the species occupancy data (MacKenzie et al., 2006). The modeling was conducted using a grid cell size of 1.5 km * 1.5 km. According to this model, the occurrence of the Siberian jay is positively associated with the cover of spruce mires (growing stock volume > 100 m^3/ha), the cover of young forests (<40 years), and length of edges between forests (growing stock volume >100 m^3/ha) and mires/fens at the territory level, and positively associated with the cover of forests (all forest types with growing stock volume > 100 m^3/ha) at the landscape level.

For autologistic and CAR models, we defined the neighborhood of a cell to encompass its eight adjacent neighbors. For all model coefficients (α, β_k and autologistic coefficient β_{k+1}) we used normally distributed non-informative priors. For variance σ_ρ^2 in CAR models we used an inverse gamma prior, $\sigma_\rho^2 \sim \mathrm{IG}(2, b_p)$ (see Latimer et al., 2006). For all three methods (non-spatial model, autologistic regression and CAR model), we run the model with two MCMC chains. Parameter estimates are based on the 20 000 iterations, after the convergence of the chains. The Bayesian analyses were conducted with OpenBugs (v.3.1.2) (Thomas et al., 2006).

In addition to comparing the parameter estimates produced by different methods (non-spatial, autologistic regression and CAR methods), we evaluated the methods on three grounds. These were: 1) Minimum Predicted Area (MPA); 2) Differences in the final predictive values for the cell occupancy; and 3) Residual autocorrelation checks based on Moran's I values.

Minimum Predicted Area (MPA) stands for the areal proportion of the total study area that excludes 10 % of known presences based on the predicted probability value for species occurrence (for the introduction of the measure, see Engler et al., 2004 and also Laita and Mönkkönen, unpublished). The exclusion is based on the threshold value for predicted probability, corresponding the probability value of the $(0.1*n)^{th}$ presence record in the ordered set of the records (where n stands for the number of known presence records). The smaller the MPA value, the smaller is the area that comprises the 90 % of the known presences. The MPA value is thus a measure of the discriminatory capacity of a model, and is especially well-founded for rare species like the Siberian Jay whose area of occupancy is expected to be quite small.

We evaluated differences in the predictions from the three modeling methods both cartographically and quantitatively. The quantitative evaluation was based on three grounds: 1) layer correlations for the occupancy predictions from the different methods (to investigate to what extent the predictions from the methods are congruent); 2) sum and average of the absolute, pairwise differences in the predictions over all 8346 grid cells (to see how similar occupancy predictions the methods produce); and 3) a check on the distribution of the pairwise differences in the occupancy predictions (to see whether some of the methods systematically over-/underestimate the occupancy predictions in relation to the others).

We calculated Moran's I values with the Spatial Analysis in Macroecology (SAM) programme (v. 4.0) (Rangel et al., 2006) to compare the alternative modeling approaches in terms of their ability to account for spatial autocorrelation in the species distribution. Distance classes were formed using the bin width of 2.5 km (approximately equaling to distance between the center of a cell and its cardinal neighbors, 2.12 km). Because the true occupancy state of the grid cell with imperfect detection was unknown, we only qualified those imperfectly detected grid cells (i.e., cells without Jay observations) whose predicted probability of occupancy was < 0.4 based on the CAR model (as also done in our previous study).

Landscape connectivity

We calculated the variable for landscape connectivity based on the predictive probability surface from the non-spatial model and CAR-model. The probability surface as predicted by the CAR model already includes the accounting for the aggregation factors, while the surface based on the non-spatial model includes only the effect of modeled environmental covariates. We were interested to see which one of surfaces forms the best basis for connectivity evaluation (as evaluated by the subsequent model fit), and whether the connectivity evaluation carried out on non-spatially predicted surface and further coupled with non-spatial modeling may provide a competing alternative to CAR-approach. The cells with a probability value greater than 0.7 were considered as source habitat cells contributing to landscape connectivity. We constructed the patch structure for connectivity calculation using a four-neighbor rule, and removed all single cell patches. Exclusion of single cell patches with the size corresponding to that of an average individual territory both speeded up the analyses as well as focused the analysis on the large-scale pattern of connectivity. Especially the non-spatial model yields many single-cell, scattered patches that are not very potential candidates for species presence due to the spatial constraints.

We calculated cost-modified distances between all pair of patches with Spatially Explicit Landscape Simulator (SELES) programme v. 3.4. (Fall and Fall, 2001). Forested cells with growing stock volume greater than 100 m^3/ha were assigned a cost distance value of 1, and all other cells a cost distance value of 5. The original cost layer of the resolution of 25 m*25 m was resampled to the resolution of 100 m *100 m with the majority rule. The Siberian Jay is exposed to predation in open areas. Although the open areas do not represent total dispersal barriers to the species, it has been documented to be unwilling to cross vast open areas in a landscape (von Haartman et al., 1963-1972).

We used negative exponential dispersal kernel to calculate the probability of dispersal between patches, and set a dispersal distance of 5 km to correspond to dispersal probability of 0.5. According to the study by Gienapp and Merilä (2011), the average dispersal distance of the species is ca. 5 km (females=5.9 km, males=5.0 km). We calculated the connectivity values for our study sites (i.e., the cells that were sampled for Siberian jay occurrence) with Conefor Sensinode v. 2.2 (Saura and Torné, 2009) using so called node removal analysis. In the node removal analysis, those study sites that were outside of the habitat patches were in turn added to the existing network, while the study sites included to the network were removed from the network. The induced change in network connectivity from adding and removing sites represents the connectivity contribution of the site (Keitt et al., 1997). As a connectivity index we used area-weighted flux, AWF, which is calculated with the following formula:

$$\mathrm{AWF} = \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} p_{ij} \cdot a_i \cdot a_j$$

,where n is the total number of patches, p_{ij} is the probability of direct dispersal between patches i and j , and a_i and a_j are the sizes of patches i and j. Because the study sites were all of equal size (1.5 km* 1.5 km), the connectivity values tell about their position in a network in relation to source habitat patches. The connectivity value of a site can be interpreted as a relative flux of dispersing individuals entering the site, assumed that the flux is controlled by the size of the donating patch and the effective distance separating the two patches. The connectivity value of a study site rises with increasing amount and closeness of habitat in the surrounding landscape (see Laita et al. 2011).

For the CAR modeling, which needs continuous surface of predictor values, we used inverse distance weighted interpolation method to extract the connectivity information from the 563 study sites to the whole study area (Fig. 1). Calculation of the connectivity values separately for all grid cells would be computationally too heavy procedure, so the interpolation offered a lightened way to end up with the connectivity estimates for the whole study area. The connectivity value of an interpolated cell was determined based on the values of its ten closest study sites, with weight given to the sites decreasing inversely to distance raised to the power of two. The interpolated surface passed through the original values of the study sites.

We fed the interpolated surface as a new variable to a subsequent modeling. We produced both non-spatial and CAR models, incorporating this new variable along with the original environmental covariates, to inspect whether this connectivity approach could either supplement the spatial approach offered by CAR, or serve as a competing alternative when coupled with the non-spatial modeling. The model evaluation was conducted based on three criteria: 1) model fit in terms of model deviance: 2) MPA value; and 3) capture of residual autocorrelation based on Moran's I values.

RESULTS

The parameter estimates from the non-spatial and CAR models were quite similar in magnitude, whereas autologistic regression produced systematically lower parameter estimates for all environmental covariates (Fig. 2). The only parameter for which the autologistic and CAR model provided a similar estimate, differing from the one provided by the non-spatial model, was the detection probability. While the non-spatial and CAR models valuated all the model parameters as statistically significant (i.e., their 95 % confidence interval does not cover zero), the autologistic model identified only two parameters for environmental covariates as significant (Fig. 2). This is contrasted with a statistically significant, high term for the autologistic covariate (posterior mean=5.0, the 95th confidence interval [3.0, 6.9]). The most notable difference in parameter estimate and its significance between the autologistic regression and the CAR methods concerned the landscape level forest cover variable. This variable possessed the highest degree of spatial structure of the variables included to the model. Confidence intervals for the parameter estimates were somewhat wider for the two spatial methods than for the non-spatial model (Fig. 2).

The autologistic regression model produced strikingly more aggregated predictive surface for the species occupancy than did the non-spatial and CAR-model (Fig. 3). The degree of aggregation produced by the CAR model was intermediate with respect to the predicted aggregation from the non-spatial and autologistic regression model. The performance of the two spatial methods, evaluated based on the MPA-value, was roughly at the same level (CAR model = 19.7 %,; autologistic regression = 24.5 %, somewhat in favor of the CAR model). The two spatial methods clearly beat the non-spatial model in model performance. For example, the MPA value of the non-spatial model was much higher than that of spatial models (Table 1). The autologistic regression seemed to capture residual spatial autocorrelation most efficiently of the modeling methods compared, but its superiority to CAR was not particularly striking (Fig. 4). In all the three methods, the residual autocorrelation was significant up to two distance classes (i.e., 5 km) (based on 999 permutations, and progessive Bonferroni correction carried out using the method represented by Legendre and Legendre (1998)).

The whole layer correlations for the occupancy predictions (posterior means) were 0.918 for the non-spatial model vs. the CAR model; 0.895 for the non-spatial vs. the autologistic model; and 0.869 for the CAR model vs. the autologistic model, suggesting that the model predictions were rather well congruent. The summed absolute differences in the predictions (posterior means) of the methods were 599.6 for the non-spatial vs. the CAR model; 655.6 for the non-spatial vs. the autologistic model; and 1255.1 for the CAR model vs. the autologistic model (yielding an average cell-wise difference in the predicted occurrence probability of 0.07, 0.08 and 0.15, respectively). Thus, while the two spatially explicit methods seemed to outperform the non-spatial methods in model performance, the predictions from the two spatial methods were more inconsistent than they were in comparison with the non-spatial model. The distributions for the pairwise differences in model predictions show that differences were roughly normally distributed and centered on zero indicating that any of the methods did not produce systematically lower/higher predictions than the others (Fig. 5). The grid cells with the greatest differences between the methods did not seem to coincide with the spatial patterning in the sampling (compare the locations of the study sites in Fig. 1). Still, the two spatial methods identified a similar set of hot spot areas for species occupancy, while those aggregated areas did not stand out so clearly from the predictions of the non-spatial model. So, while the predictions from the spatial methods were not completely matching at the cell level, they still captured the same broad distributions patterns.

It clearly seemed that the predicted occupancy pattern produced by the CAR model (CAR-connectivity) offers a better basis for connectivity evaluation than the one produced by the non-spatial model (non-spatial connectivity; Table 1). Although all our three evaluation criteria showed that the inclusion of CAR-based landscape connectivity as an additional explanatory variable led to improvement in model fit (lower deviance), performance (slightly lower MPA) and the capture of residual autocorrelation (lower Moran's I) in the non-spatial models, the model even after inclusion of landscape connectivity could not compete with the CAR model without connectivity (Table 1). In the CAR models, the incorporation of the CAR-based landscape connectivity only slightly increased model fit and reduced residual autocorrelation, but it was associated with a considerable increase in the MPA value. So, the evaluation criteria did not unequivocally support the modeling benefits associated with the coupling of landscape connectivity component with the CAR approach.

DISCUSSION

The CAR approach and autologistic regression produced rather differing coefficient estimates for the environmental covariates; the means of the posterior distributions for the coefficients were systematically greater for the CAR model than for the autologistic regression. Carl and Kühn (2007) and Dormann (2009) have criticized the autologistic regression for its circularity; because the covariate is calculated as the mean of the predicted values in the grid neighborhood, the autocovariate is thus confusing the response with environmental variables. The autocovariate can thus overdominate parameter estimations, deflating the importance of environmental covariates. This has also been confirmed in several studies that have also showed that this underestimation can be most severe for the highly spatially structured environmental covariates (e.g., Betts, et al. 2009). One of the territory level covariate and the landscape level covariate turned out to lose their significance in our autologistic regression model (compared to nonspatial and CAR models). This makes sense considering the strong spatial structure of the landscape level variable, but our result point out also a general tendency for autologistic regression models to underestimate the parameter values. Based on autologistic modeling, it is more difficult to come up with convincing evidence for the importance of landscape level covariates affecting the probability of species presence than it is for variables with weaker spatial autocorrelation (due to the shared scale of variation between the landscape level covariate and the response). On the other hand, non-spatial models tend to exaggerate the importance of the variables with a strong spatial structure (see e.g. Lennon, 2000; Lichstein et al., 2002; Segurado et al., 2006).

It has been suggested that the autologistic regression may produce more correct parameter estimates in the case of a notable share of missing response values rather than in the case of almost total coverage for the response (Dormann, 2009). The missing data may help the autologistic regression in parameter estimation so that its allocation between space and environment is tuned towards the correct environmental emphasis. However, there is also a lower limit for the safe share of missing response values. In the study of Latimer et al. (2006), the parameter estimates were stable down to the subsample of 5% of the full coverage data. We had altogether occupancy data for 6.7 % of all the grid cells. Thus, this share should be above the safe limits in terms of parameter estimations. On the other hand, for such a great share of missing values the parameter estimates produced by the autologistic regression should not be severely biased in favor of the space allocation.

As opposed to autologistic regression, CAR models do not represent the autocorrelation effect in the form of an additional variable. CAR models incorporate a spatially correlated error to the random component of the model, which may help to keep the effect of space and environment apart. Our results showed that the CAR model is inevitably better in detecting the signal from the environmental factors than the autologistic regression, but with the field data on species distribution such as ours it is difficult to conclude which is the true relative magnitude of environmental vs. spatial effects. The indisputable interrelation of environment and space could be further break apart by hierarchical partitioning methods (see e.g., Cushman and McGarigal, 2002). However, for variables with a strong spatial structure, the variation is mainly and unavoidably captured by the partition for the shared variation between the environment and space. So, even the partitioning methods may provide only a limited solution to the problem of highly spatially structured variables in controlling species distributions.

If the modeling efforts aim at prediction instead of inference, a possible biased allocation between environmental covariates and space is not necessarily harmful. Even though autologistic regression underestimated the contribution of environmental covariates, it still performed better than the non-spatial model. However, our results show that this enhanced model performance of the spatial models does not come in one consistent form, but in terms of the predictions for cell occupancy, the two spatial models differed more from each other than either spatial model from the non-spatial model. Without more field data on species occupancy, it is impossible to rank the two spatial methods with respect to their ability to correctly predict the species distribution.

Our results show that it is not so straightforward to evaluate the outcomes of different modeling approaches; the approaches may have different ranks depending on the criteria based on which they are evaluated. For example, while the autologistic regression seem to be more efficient than CAR modeling in removing residual autocorrelation, the CAR modeling has a higher discriminatory capacity (based on MPA) than the autologistic regression. The cartographic representation of the model predictions (Fig. 3) contains much information that was hardly captured by the quantitative evaluation criteria. The predictive surfaces of model predictions showed efficiently the different tendency of the methods to aggregate species occupancy (in the order of decreasing aggregation; the autologistic, the CAR and the nonspatial model). The ultimate goal of modeling should determine the evaluation criteria based on which the modeling outcomes are compared.

We have compared the performance of the methods within one area, and, due to the high computational demands of the spatial Bayesian analysis for a large study area, used all the data for parameter estimation instead of splitting it up (or resubstituting) for within-area validation. As a measure of model performance, we have used the degree to which a given model can explain and predict the partly known recent distribution pattern (and at the same time, avoiding overprediction). So, on these grounds, it is impossible to conclude which one of the two spatial approaches has greater potential to generalize and transfer to new areas. Ability to generalize model results within the area and among areas (i.e., model transferability to new areas) are two different modeling aspirations, and we do not know whether they are simultaneously achievable or in conflict (but see e.g., Heikkinen et al., 2011). It might be that that some of the enhanced model performance of the spatial models (when compared to non-spatial one) comes with the risks of overcompensation for the space effect as found in this study and by Segurado et al. (2006) for the autologistic regression model. This possible overcompensation of autologistic models with an associated underestimation of environmental control lowers model transferability (see e.g., Carroll et al., 2010). Because the alternative spatial approaches came up with different inference concerning the effect of environmental covariates in controlling species distribution, it is likely that the methods may show differences in their within and among-area generalisability. The relative merits and transferability of model outcomes from different spatial modeling methods are interesting topics for future research.

According to our results, CAR approach was superior to the non-spatial approach even when complemented with the landscape connectivity component. This holds true even though the landscape connectivity component enhanced the performance of the non-spatial model. The difference in model fit and performance between the CAR approach and the non-spatial model with the connectivity component can be considered a clue to spatial signature produced by extrinsic factors. This is because spatial component of the CAR model incorporates autocorrelation produced both by extrinsic and intrinsic factors, while the landscape connectivity component emphasizes intrinsic factors. The difference in favor of the CAR approach, compared to non-spatial approach with the connectivity component, may have been even greater if we would have screened for an optimum neighborhood size (instead of the fixed second-order neighborhood).

We used predicted probability values for species occurrence (represented by posterior sample means) as a fixed baseline for the evaluation of landscape connectivity, and fed this so-formed new covariate to the second round of the modeling. In an ideal case, the connectivity covariate and its associated parameter would also be evaluated in an iterative manner along with all the other model parameters. This approach would acknowledge the fact that the connectivity component is related to the probability of species occurrence, but it would come with a huge computational load. We think that the connectivity evaluations based on our two-step procedure, in addition to analytical simplicity, serves as a good approximation to the iteratively achieved connectivity evaluations.

Our results show that space and environment are strongly interrelated (see also Currie, 2007). This causes problems especially for parameter estimation and inference so that different modeling methods have different sensitivities to allocate explanatory effect between environment and space. However, our results show that the methods may still achieve the same level of model performance via different allocations. In this study, we also tested the possibility of using a measure of landscape connectivity as an alternative and/or supplementary way (compared with the normally used neighborhood based connectivity measures) of handling the spatially structured dispersal process in distribution modeling. Even though the use of landscape connectivity clearly lacked behind the normal CAR approach in the modeling outcome, it may still provide insights to the nature and strength of intrinsic factors underlying species distributions. Although the different spatial methods do not obviously provide any standard way to cope with the effects of autocorrelation, they allow important insights into the controlling factors of species distributions in spatially structured environments where the ecological processes themselves are limited by distance. Species distribution modeling may gain in the form of enhanced predictive power and more profound understanding of distribution patterns than on pure non-spatial grounds. Therefore, species distribution modeling has a potential to provide a more solid base for species conservation.

ACKNOWLEDGEMENT

The study was funded by Kone Foundation and the rector of University of Jyväskylä (to AL). We thank the ornithological societies, Central Finland Bird Club, Lintuyhdistys Kuikka, North Karelian Ornithological Society and Suomenselän Lintutieteellinen yhdistys ry, and Metsähallitus for the data on the Siberian Jay occurrences, and the Finnish Forest Research Institute for the multi-source NFI-data.

REFERENCES

With K.A. & King A.W. 1999. Extinction thresholds for species in fractal landscapes. Conserv. Biol. 13, 314-326.

Araújo, M.B. and Williams, P.H., 2000. Selecting areas for species persistence using occurrence data. Biol.Conserv. 96, 331-345.

Augustin, N.H., Mugglestone, M.A. and Buckland, S.T., 1996. An autologistic model for the spatial distribution of wildlife. J.Appl.Ecol. 33, pp. 339-347.

Augustin, N.H., Mugglestone, M.A. and Buckland, S.T., 1998. The role of simulation in modeling spatially correlated data. Environmetrics. 9, 175-196.

Bauerfeind, S., Theisen, A. and Fischer, K., 2009. Patch occupancy in the endangered butterfly *Lycaena helle* in a fragmented landscape: effects of habitat quality, patch size and isolation. J.Insect Conserv. 13, 271-277.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. and Elston, D.A., 2010. Regression analysis of spatial data. Ecol.Lett. 13, 246-264.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J.Roy. Statistical Society, Series B. 36, 192-236.

Betts, M.G., Ganio, L.M., Huso, M.M.P., Som, N.A., Huettmann, F., Bowman, J. and Wintle, B.A., 2009. Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review" Ecography. 32, 374-378.

Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V.B., Akre, T.S.B., Albaladejo, R.G., Albuquerque, F.S., Aparicio, A., Araújo, M.B., Baselga, A., Beck, J., Bellocq, M.I., Böhning-Gaese, K., Borges, P.A.V., Castro-Parga, I., Chey, V.K., Chown, S.L., Marco, P.d., Jr, Dobkin, D.S., Ferrer-Castán, D., Field, R., Filloy, J., Fleishman, E., Gómez, J.F., Hortal, J., Iverson, J.B., Kerr, J.T., Kissling, W.D., Kitching, I.J., León-Cortés, J.L., Lobo, J.M., Montoya, D., Morales-Castilla, I., Moreno, J.C., Oberdorff, T., Olalla-Tárraga, M.Á., Pausas, J.G., Qian, H., Rahbek, C., Rodríguez, M.Á., Rueda, M., Ruggiero, A., Sackmann, P., Sanders, N.J., Terribile, L.C., Vetaas, O.R. and Hawkins, B.A., 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. Ecography. 32, 193-204.

Brouwers, N. and Newton, A., 2009. The influence of habitat availability and landscape structure on the distribution of wood cricket (Nemobius sylvestris) on the Isle of Wight, UK. Landscape Ecol. 24, 199-212.

Brown, J.H. and Kodric-Brown, A., 1977. Turnover rates in insular biogeography: effect of immigration on extinction. Ecology. 58, 445-449.

Carl, G. and Kühn, I., 2007. Analyzing spatial autocorrelation in species distributions using Gaussian and logit models. Ecol.Model. 207, 159-170.

Carroll, C., Johnson, D.S., Dunk, J.R. and Zielinski, W.J., 2010. Hierarchical Bayesian spatial models for multispecies conservation planning and monitoring. Conserv.Biol. 24, 1538-1548.

Cressie, N.A.C., 1993. Statistics for spatial data. Wiley, New York.

Currie, D.J., 2007. Disentangling the roles of environment and space in ecology. J.Biogeogr. 34, 2009-2011.

Cushman, S.A. and McGarigal, K., 2002. Hierarchical, multi-scale decomposition of species-environment relationships. Landscape Ecol. 17, 637-646.

Dormann, C.F., 2007. Assessing the validity of autologistic regression. Ecol.Model. 207, 234-242.

Dormann, C.F., 2009. Response to Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review" Ecography. 32, 379-381.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kuhn, I., Ohlemuller, R., Peres-Neto, P.R., Reineking, B., Schroder, B., Schurr, F.M. and Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography. 30, 609-628.

Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M. and Essl, F., 2009. Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. Biol.Invasions. 11, 2401-2414.

Engler, R., Guisan, A. and Rechsteiner, L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. J.Appl.Ecol. 41, 263-274.

Fall, A. and Fall, J., 2001. A domain-specific language for models of landscape dynamics. Ecol.Model. 141, 1-18.

Farnsworth, M.L., Hoeting, J.A., Hobbs, N.T. and Miller, M.W., 2006. Linking chronic wasting disease to mule deer movement dcales: A hierarchical Bayesian approach. Ecol.Appl. 16, 1026-1036.

Gienapp, P. and Merilä, J., 2011. Sex-specific fitness consequences of dispersal in Siberian jays. Behav.Ecol.Sociobiol. 65, 131-140.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R. and Hastie, T., 2006. Making better biogeographical predictions of species' distributions. J.Appl.Ecol. 43, 386-392.

Guisan, A. and Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol.Model. 135, 147-186.

Haartman, von L., Hilden, O., Linkola, P., Suomalainen, P. and Tenovuo, R., 1963-1972. Pohjolan linnut värikuvin II. Otava, pp. 445-1092.

Heikkinen, R.K., Marmion, M. and Luoto, M., 2011. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography (in press).

Hoeting, J.A., 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. Ecol.Appl. 19, 574-577.

Hoeting, J., Leecaster, M. and Bowden, D., 2000. An improved model for spatially correlated binary responses. J.Agric.Biol.Envir.S. 5, 102-114.

Huffer, F.W. and Wu, H., 1998. Markov Chain Monte Carlo for autologistic regression models with application to the distribution of plant species. Biometrics. 54, 509-524.

Ishihama, F., Takeda, T., Oguma, H. and Takenaka, A., 2010. Comparison of effects of spatial autocorrelation on distribution predictions of four rare plant species in the Watarase wetland. Ecol.Res. 25, 1057-1069.

Keitt, T.H., Urban, D.L. and Milne, B.T., 1997. Detecting critical scales in fragmented landscapes. Conserv.Ecol. 1, 4.

Kissling, W.D. and Carl, G., 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. Global Ecol.Biogeogr. 17, 59-71.

Kuhn, I., 2007. Incorporating spatial autocorrelation may invert observed patterns. Divers.Distrib. 13, 66-69.

Laita, A., and Mönkkönen, M. (unpublished). Space, scales and imperfect detection – hierarchical modeling of the habitat association of the Siberian Jay (*Perisoreus infaustus)* in boreal forest landscapes.

Laita, A., Kotiaho, J.S. and Mönkkönen, M., 2011. Graph-theoretic connectivity measures - what do they tell us about connectivity? Landscape Ecol. 26, 951-967.

Latimer, A.M., Wu, S., Gelfand, A.E. and Silander Jr., J.A., 2006. Building statistical models to analyze species distributions. Ecol.Appl. 16, 33-50.

Legendre, P., 1993. Spatial Autocorrelation - trouble or new paradigm. Ecology. 74, 1659-1673.

Legendre, P. and Legendre, L., 1998. Numerical ecology. Elsevier, Amsterdam.

Lennon, J.J., 2000. Red-shifts and red herrings in geographical ecology. Ecography. 23, 101-113.

Lichstein, J.W., Simons, T.R., Shriner, S.A. and Franzreb, K.E., 2002. Spatial autocorrelation and autoregressive models in ecology. Ecol.Monogr. 72, 445-463.

MacKenzie, D.I., 2006. Occupancy estimation and modeling : inferring patterns and dynamics of species. Elsevier Academic Press, Burlington.

Miller, J., Franklin, J. and Aspinall, R., 2007. Incorporating spatial dependence in predictive vegetation models. Ecol.Model. 202, 225-242.

Osborne, P.E., Alonso, J.C. and Bryant, R.G., 2001. Modeling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. J.Appl.Ecol. 38, 458-471.

Piorecky, M.D. and Prescott, D.R.C., 2006. Multiple spatial scale logistic and autologistic habitat selection models for northern pygmy owls, along the eastern slopes of Alberta's Rocky Mountains. Biol.Conserv. 129, 360-371.

Rangel,Thiago Fernando L. V. B., Diniz-Filho, J.A.F. and Bini, L.M., 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. Global Ecol.Biogeogr. 15, 321-327.

Ranius, T., Johansson, V. and Fahrig, L., 2010. A comparison of patch connectivity measures using data on invertebrates in hollow oaks. Ecography. 33, 971-978.

Royle, J.A. and Dorazio, R.M., 2008. Hierarchical modeling and inference in ecology : the analysis of data from populations, metapopulations and communities. Academic Press, San Diego.

Saura, S. and Torné, J., 2009. Conefor Sensinode 2.2: A software package for quantifying the importance of habitat patches for landscape connectivity. Environmental Modeling & Software. 24, 135-139.

Segurado, P., Araújo, M.B. and Kunin, W.E., 2006. Consequences of spatial autocorrelation for niche-based models. J.Appl.Ecol. 43, 433-444.

Thomas, A., O'Hara, B., Ligges, U, and Sturtz, U. 2006. Making BUGS Open. - R News 6: 12-17.

Wintle, B.A. and Bardos, D.C., 2006. Modeling species-habitat relationships with spatially autocorrelated observation data. Ecol.Appl. 16, 1945-1958.

Wu, H. and Huffer, F.R., 1997. Modeling the distribution of plant species using the autologistic regression model. Environ.Ecol.Stat. 4, 31-48.

TABLE 1 Comparisons of non-spatial and CAR modeling methods, either including the landscape connectivity component or not, in terms of model deviance, MPA and the Moran's I value of the first distance class. The connectivity component was defined based on AWF-measure (see text), and was based either on the occupancy pattern predicted by the non-spatial model (NonSpat-connectivity) or the CAR model (CAR-connectivity).

Model	Deviance	MPA (%)	Moran's I
NonSpat	945.3	36.9	0.279
NonSpat + CAR-connectivity	896.9	36.7	0.165
NonSpat + NonSpat-connectivity	946.4	37.6	0.281
CAR	867.0	19.7	0.242
CAR + CAR-connectivity	861.1	28.9	0.168
CAR + NonSpat-connectivity	867.3	20.4	0.252

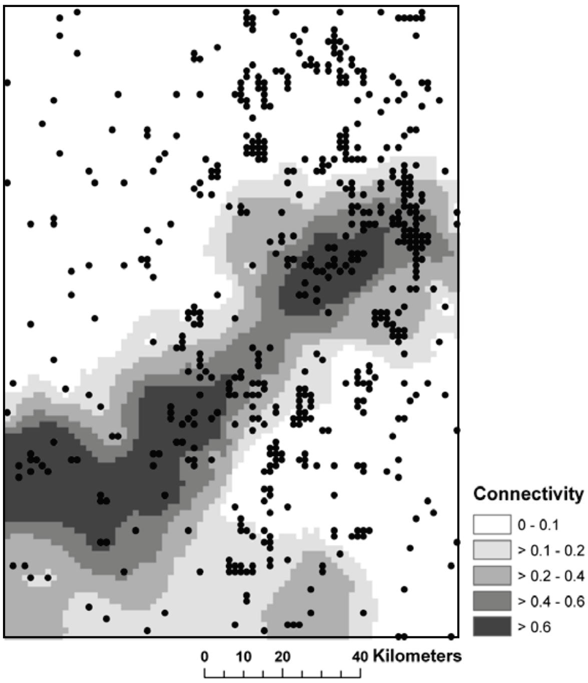


FIGURE 1 Interpolated surface for the grid connectivity, i.e. relative measure for the degree of which a grid cell is connected to the source cells, based on the occupancy pattern predicted by the CAR-model. Connectivity is measured based on area-weighted flux – measure. Black points represent the study sites (i.e., grid cells with sampling, represented as centers of the cells) whose connectivity information is interpolated for the whole area using inverse distance weighted method. This interpolated layer is used as a new covariate in subsequent spatial modelling.

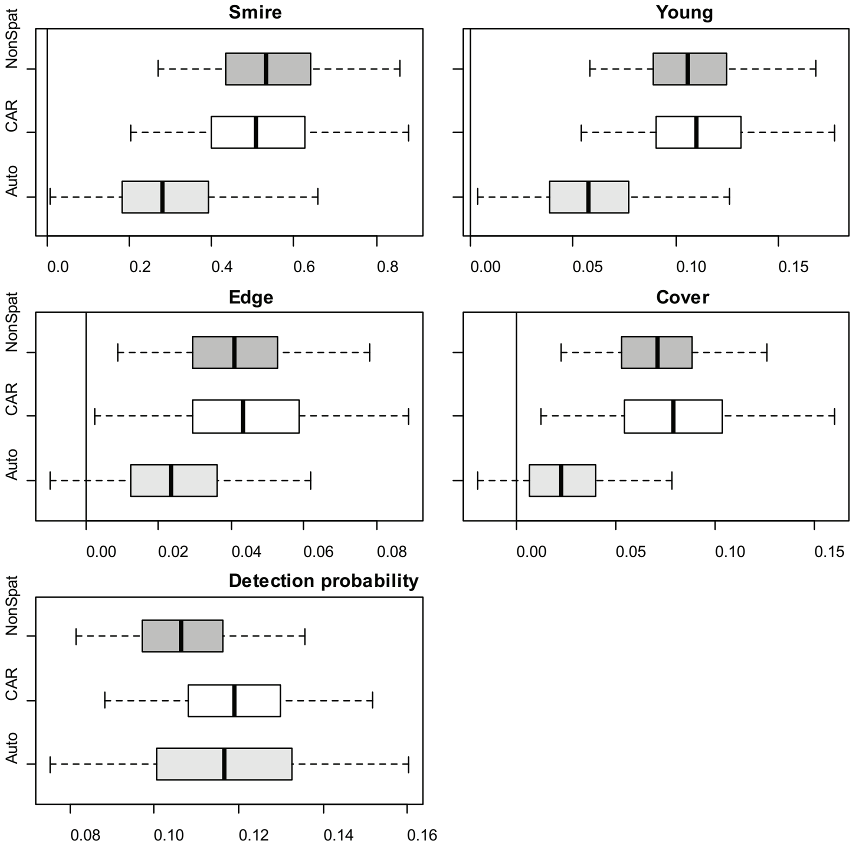


FIGURE 2 Posterior samples for the parameters based on three modeling approaches; non-spatial, autologistic regression and CAR. The posterior samples represent two-chain runs, both thinned by ten. In total, both thinned chains comprise 2000 iterations, sampled after the chain convergence. The middle line shows the median, bars the quartiles (25% and 75 %) of the distribution and whiskers the 2.5th and 97.5th percentiles of the distribution (i.e. they show the 95 % confidence interval for the parameter). *Smire* and *Young* stands for the cover of spruce mires (> 100m^3/ha) and young forests (<40 years) at the territory level, respectively; *Edge* for the length of edges shared between forests and mire/fens at the territory level; *Cover* for the cover of forest (>100 m^3/ha) at the landscape level; and *Detection probability* for the probability of seeing a Siberian Jay in a visit to an occupied site.

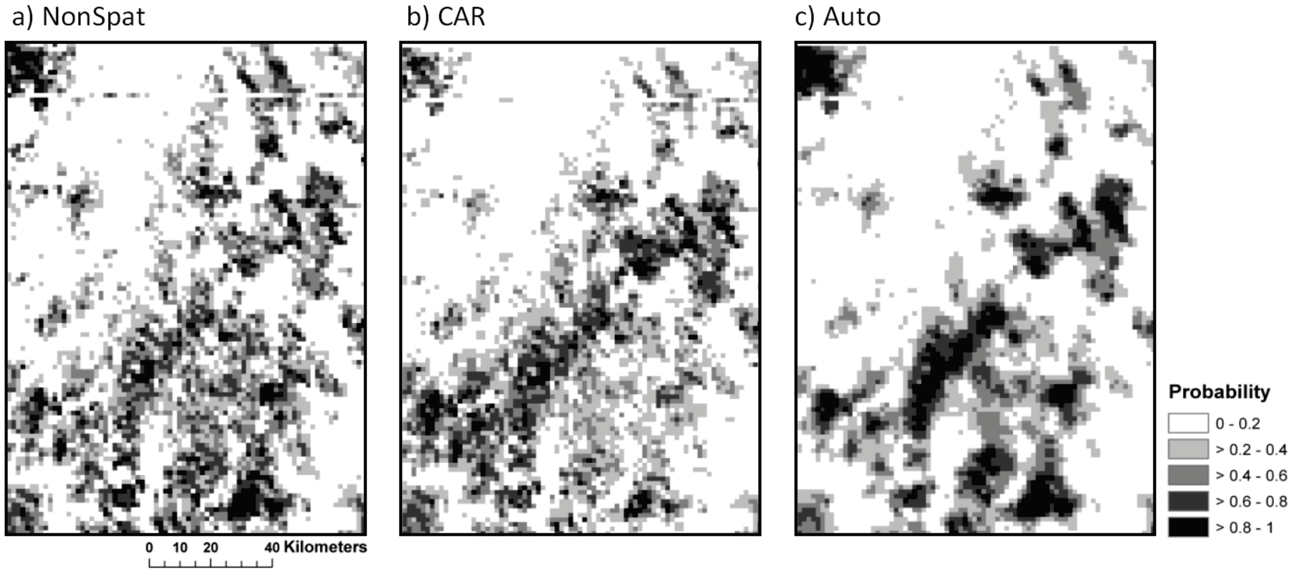


FIGURE 3 The prediction for the species presence based on the Bayesian non-spatial logistic regression, CAR model and the autologistic regression model. The sampling covers 6.7 % of the study area (see the cells with sampling in Fig. 1).

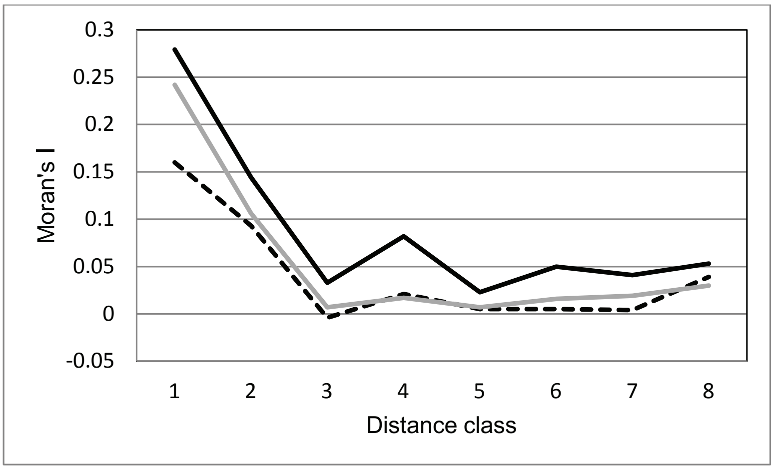


FIGURE 4 Moran’s I values of the first eight distance classes (bin width of 2.5 km, ranging up to 20 km) for the three modeling methods; nonspatial (black line), CAR (grey line) and autologistic regression (dashed line).

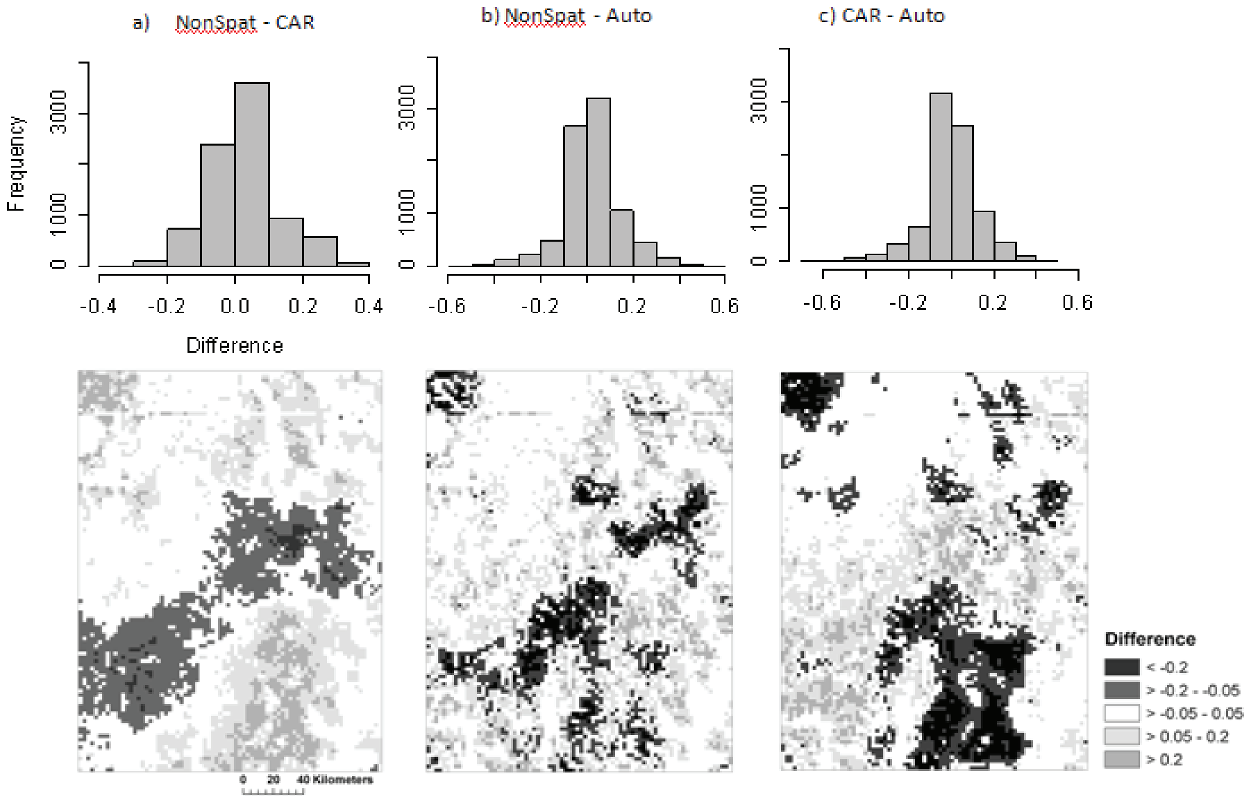


FIGURE 5 The distributions for the differences in the predictions(represented as posterior mean values) of the models for a) the non-spatial model minus the CAR model; b) the non-spatial model minus the autologistic regression; c) CAR model minus the autologistic regression.