



Ecology Research Handbook

Volume IV

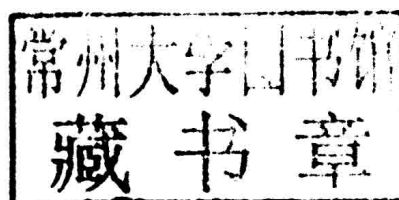
Liam Page



Ecology Research Handbook

Volume IV

Edited by **Liam Page**



New York

Published by Callisto Reference,
106 Park Avenue, Suite 200,
New York, NY 10016, USA
www.callistoreference.com

Ecology Research Handbook: Volume IV
Edited by Liam Page

© 2015 Callisto Reference

International Standard Book Number: 978-1-63239-162-9 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Copyright for all individual chapters remain with the respective authors as indicated. A wide variety of references are listed. Permission and sources are indicated; for detailed attributions, please refer to the permissions page. Reasonable efforts have been made to publish reliable data and information, but the authors, editors and publisher cannot assume any responsibility for the validity of all materials or the consequences of their use.

The publisher's policy is to use permanent paper from mills that operate a sustainable forestry policy. Furthermore, the publisher ensures that the text paper and cover boards used have met acceptable environmental accreditation standards.

Trademark Notice: Registered trademark of products or corporate names are used only for explanation and identification without intent to infringe.

Printed in China.

Ecology Research Handbook

Volume IV

Preface

The origins of the word “ecology” can be traced to the year 1866, when it was coined by the German scientist Ernst Haeckel. Ecology is an interdisciplinary field, which is the scientific study of interactions among organisms and their environment. The foundations of this subject were laid by Ancient Greek philosophers such as Hippocrates and Aristotle, who studied natural history. From the 19th century onwards, concepts such as natural selection and adaptation transformed Ecology into a more rigorous discipline.

The concepts of Ecology are premised on ecosystems, which include organisms, the communities they make up, and the non-living components of their environment. The scope of Ecology spans a wide array of interacting levels of organisms, ranging from the micro-level to planetary scale. This subject helps comprehend how the living world interacts. It provides evidence on the interdependence between the inanimate and animate elements of the ecosystem.

Ecology also covers topics of subjects like Biology and Earth Science. Topics of interest to ecologists include diversity, distribution, and population of organisms. The manner in which biodiversity affects ecological function is an important and emerging focus area in ecological studies. A better understanding of ecological systems is crucial in contemporary times, as it allows scientists to predict the consequences of human activity on the environment.

I would like to thank our researchers and writers for sharing their valuable research with us in this book.

Editor

Contents

| | | |
|------------|---|------------|
| | Preface | VII |
| Chapter 1 | Landscape spatial structure for predicting suitable habitat: The case of <i>Dalea villosa</i> in Saskatchewan S. Lowe, X. Guo, D. Henderson | 1 |
| Chapter 2 | Factors associated with the spread of Chinese Tallow in East Texas forestlands Zhaofei Fan, Yuan Tan, Michael K. Crosby | 15 |
| Chapter 3 | Salvage logging versus natural regeneration post-fire practices in a forest: Soil chemical and microbial aspects Orit Ginzburg, Yosef Steinberger | 25 |
| Chapter 4 | Seed dispersal by Tana River mangabeys in fragmented gallery forests Duncan M. Kimuyu, Geoffrey M. Wahungu, Dennis O. Otieno | 34 |
| Chapter 5 | A chemical ecological investigation of the allelopathic potential of <i>Lamium amplexicaule</i> and <i>Lamium purpureum</i> Chelsea D. Jones, Katherine E. Woods, William N. Setzer | 43 |
| Chapter 6 | Insect herbivory along environmental gradients Nigel R. Andrew, Isobel R. Roberts, Sarah J. Hill | 54 |
| Chapter 7 | Effects of micro-environmental conditions and forest disturbance on the establishment of two Andean palms in Ecuador Diana Rodríguez-Paredes, Rommel Montúfar-Galárraga, Henrik Meilby | 66 |
| Chapter 8 | Robots in Ecology: Welcome to the machine David Grémillet, William Puech, Véronique Garçon, Thierry Boulinier, Yvon Le Maho | 77 |
| Chapter 9 | Photosynthetic toxicity and oxidative damage induced by nano-Fe₃O₄ on <i>Chlorella vulgaris</i> in aquatic environment Xiaoxiao Chen, Xing Zhu, Rui Li, Hanchao Yao, Zhisong Lu, Xu Yang | 86 |
| Chapter 10 | Effects of tree sapling diversity and nutrient addition on herb-layer invasibility in communities of subtropical species Sabine Both, Teng Fang, Martin Baruffol, Bernhard Schmid, Helge Bruelheide, Alexandra Erfmeier | 94 |
| Chapter 11 | Changes in soil free-living nematode communities and their trophic composition along a climatic gradient Tal Levi, Chen Sherman, Stanislav Pen-Mouratov, Yosef Steinberger | 105 |

| | | |
|------------|--|-----|
| Chapter 12 | Water quality and water-use conflicts in Lake Taabo (Ivory Coast) Noël Grogga, Allassane Ouattara, Sébastino Da Costa, Alain Dauta, Olivier Beauchard, Jacques Moreau, Germain Gourene, Pascal Laffaille | 116 |
| Chapter 13 | Spatial pattern of landscape changes and consequence changes in species diversity between 1956-1999 of <i>Pinus halepensis</i> Miller plantations in Montes de Malaga State Park (Andalusia, Spain) Rafael Maria Navarro-Cerrillo, José Ramón Guzmán-Álvarez, Inmaculada Clavero-Rumbao, Carlos Ceaceros | 126 |
| Chapter 14 | Floristic composition of the halophilic and salt-resistant plant population in Hammam-Boughrara (Oran-Algeria) Amel Taleb-Bendiab Sari-Ali, Noury Benabadji, Mohamed Bouazza | 138 |
| Chapter 15 | Litter cover of the invasive tree <i>Melaleuca quinquenervia</i> influences seedling emergence and survival Min B. Rayamajhi, Paul D. Pratt, Philip W. Tipping, Ted D. Center | 151 |
| Chapter 16 | Polyplacophora species richness, composition and distribution of its community associated with the intertidal rocky substrate in the marine priority region No. 32 in Guerrero, Mexico Rafael Flores-Garza, Lizeth Galeana-Rebolledo, Adriana Reyes-Gómez, Sergio García Ibáñez, Carmina Torreblanca-Ramírez, Pedro Flores-Rodríguez y Arcadio Valdés González | 161 |
| Chapter 17 | Exploring the edge of a natural disaster Michael W. Jenkins, Dan J. Krofcheck, Rachel Teasdale, James Houppis, James Pushnik | 171 |
| Chapter 18 | Chasing the queens of the alien predator of honeybees: A water drop in the invasiveness ocean Karine Monceau, Olivier Bonnard, Denis Thiéry | 182 |
| Chapter 19 | The current state of the plant diversity in the Tlemcen region (Northwest Algeria) Smaïn El-Amine Henaoui, Mohammed Bouazza | 191 |

Permissions

List of Contributors

Landscape spatial structure for predicting suitable habitat: The case of *Dalea villosa* in Saskatchewan

S. Lowe^{1*}, X. Guo¹, D. Henderson²

¹Department of Geography and Planning, University of Saskatchewan, Saskatoon, Canada;

*Corresponding Author

²Environmental Stewardship Branch, Prairie and Northern Canadian Wildlife Service, Environment Canada, Saskatoon, Canada

ABSTRACT

Prediction of potentially suitable habitat is important for the recovery of species protected by federal laws. Therefore, the objective of this research was to study the relationship between habitat configuration and hairy prairie-clover occurrence in order to predict suitable and unsuitable bare sand habitat across the study site. Bare sand patches were extracted from a land cover classification of the study site and several patch scaled metrics were calculated to characterize habitat spatial structure. Binary logistic regression was used to determine which metrics were significantly correlated with hairy prairie-clover occurrences. The logistic regression equation was subsequently used to predict suitable and unsuitable bare sand habitat for hairy prairie-clover based on the probability of occupancy. Results showed that about 29% of the variation in bare sand patch occupancy could be explained by the size, shape, and degree of isolation of a sand patch as well as the amount of vegetation on a sand patch in the early growing season. Based on these variables, 18.8% of bare sand patches in the study site were predicted to be unsuitable hairy prairie-clover habitat, 45.7% were predicted to be marginally unsuitable, 32.7% were predicted to be suitable, and 2.8% were predicted to be marginally suitable.

Keywords: Spatial Structure; Habitat Pattern; Remote Sensing; Habitat Suitability; Rare Plants

1. INTRODUCTION

One challenge with predictive modelling is that many species records include presence-only data with no reliable records of absences. Popular methods such as GARP (Genetic Algorithm for Rule-set Prediction), BioClim,

and Maxent have proven accuracy for modelling when presence-only data is available for a species [1,2]. However, evaluating model performance for presence-only methods still requires the use of presence/absence data. Evaluating model performance with presence-only data limits the options for the data set and the power of statistical evaluations [1]. Therefore, where reliable presence/absence data sets are available, it is best to choose methods such as generalized linear models, generalized additive models, regression-based models, or resource selection functions [2,3]. Further, most theoretical models assume that response curves are either sigmoid or gaussian, although in real ecological systems gaussian curves may be rare due to factors such as thresholds along an environmental gradient or interspecific interactions. Therefore, regression-based habitat suitability models, such as logistic regressions, may be a better choice because they can fit curves ranging from parametric functions to less constrained shapes [3].

For plant species occupying discrete habitat niches, such as in the case of hairy prairie-clover, habitat suitability modeling can be effectively applied [4]. The ecological niche concept is based on the theory that species only thrive in discrete ranges of environmental conditions [3]. However, due to abiotic and biotic interactions with the surrounding environment a species may be constrained to occupy only a subset of its fundamental niche. This subset can be described as the realized niche. Habitat suitability modeling is an application of the ecological niche theory because it aims to predict the probability of occurrence of a species based on environmental conditions. The geographical distribution of a species generally follows three constraints: 1) the local environment allows the population to grow, 2) interactions with other local species allow the species to persist, and 3) the location is accessible given the dispersal abilities of the species [3]. Therefore, habitat suitability modeling can only reconstruct the realized niche for a species based on the environmental variables present in the locations it occupies [5].

When dealing with plant species, it is important to consider that they are immobile, exhibit a strong spatial structure through specific habitat preferences, and their dispersal is restricted [4]. Therefore, the geographical distribution of a plant species is heavily dependent on the geographical distribution of its realized niche. Thus, the above three constraints can be equated to functions of patch area, patch quality, and patch isolation, respectively [6]. The three most basic processes determining the survival of a local population patch are dispersal, establishment, and persistence [4]. It has been adequately documented that dispersal is mainly affected by patch isolation and matrix quality; establishment depends on patch size, number of patches, and patch quality; and persistence depends on patch size and patch quality [7,8]. Sensitivity to these spatial factors has been noted to be higher in rare species and habitat specialists as opposed to generalist species because they tend to have lower reproductive outputs, have shorter lived seeds, are smaller in size, and are geographically restricted [8,9].

Many habitat suitability models still fail to be successfully applied to real ecological problems because they fail to address spatial structure through assuming that all patches are equivalent in size and quality, and that all local populations are equally accessible by dispersers. Considering that all plant populations are spatially structured, habitat pattern must be incorporated into habitat suitability models to allow for an ecologically meaningful interpretation [10]. However, a distinction between habitat parameters that can be mapped and measured versus habitat parameters that are ecologically relevant must be made [11].

For hairy prairie-clover, the dominant habitat requirement appears to be an element of bare sand cover resulting in the plants confinement to geographical features such as parabolic sand dunes, stabilized blow-outs, dune depressions, and sand flats [12]. It has been identified

that factors such as climate change, vegetation encroachment, invasive species encroachment, anthropogenic sources and land use changes, and loss of natural disturbance by bison and fire have contributed to the reduction in sand dune activity and bare sand areas within the prairies [13-15]. The association of hairy prairie-clover with these types of bare sand areas makes it particularly vulnerable to changes in habitat pattern as incurred by the above factors. It has also been identified that vegetation and invasive species encroachment of bare sand and sand dune areas, loss of grazing and fire disturbance pressures, and changes in land use patterns are the main threats to the survival of hairy prairie-clover [12]. Thus, it is hypothesised that the dominant factor limiting the abundance of hairy prairie-clover on the landscape is the spatial configuration of bare sand areas in context to other landscape elements.

The purpose of this research was to analyse habitat pattern to gain an understanding of the spatial structure of bare sand habitat for hairy prairie-clover and to predict potentially suitable and unsuitable habitat for this plant. There were two specific research questions: 1) What is the spatial pattern of intra-patch and inter-patch characteristics for occupied and unoccupied bare sand habitat? and 2) What spatial configuration of bare sand habitat best explains and predicts occupancy by hairy prairie-clover?

2. METHODS

2.1. Study Site

This study took place in south central Saskatchewan within the Dundurn and Rudy Rosedale Prairie Farm Rehabilitation Administration (PFRA) Community Pastures located at 51°45'00"N and 106°45'00"W (Figure 1). The area is a remnant patch of natural mixed-grass prairie and sand dune complexes bordered on the north

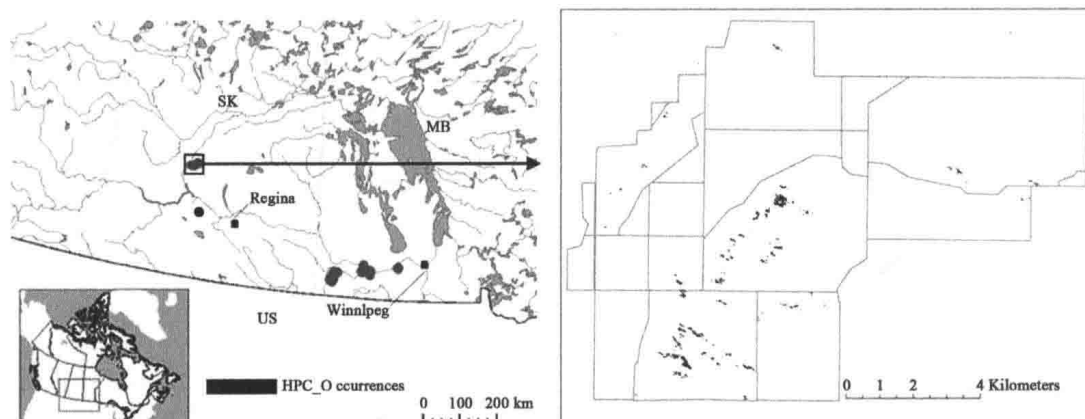


Figure 1. Dundurn and Rudy Rosedale Prairie Farm Rehabilitation Administration Community Pastures, Saskatchewan. The Canadian distribution of hairy prairie-clover (HPC) is shown on the left [20]. A close-up of the study site is shown on the right with pasture boundaries outlined in grey. The known distribution of the population is shown by the black polygons.

by the Canadian Forces Base, 17-Wing Detachment Dundurn, the east and south by mixed farm land, and the west by highway #219. The landscape is dominated by five main land cover types; aspen (*Populus tremuloides* Michx.), shrub (*Elaeagnus commutata* Bernh. ex. Rydb., *Prunus virginiana* L. var. *virginiana*, *Symphoricarpos occidentalis* Hook.), juniper (*Juniperus horizontalis* Moench.), grassland, and sand dunes. This pasture is categorized as a low dunes ecosite with mainly stabilized sandy hills that have gentle to moderate slopes and areas between dunes that are relatively flat comprising a mixed-grass community [16].

The sand dunes found in the Dundurn area comprise four physiographic categories: active complexes, stabilized blow-outs, stabilized dunes, and dune depressions. Species most commonly associated with active complexes are *Psoralea lanceolata*, *Agropyron spp.*, and *Lygodesmia juncea*. Stabilized blow-outs are saucer-shaped depressions oriented in the direction of the prevailing wind and are most often vegetated by *Juniperus horizontalis*, *Carex heliophila* and *Selaginella densa*. When dunes become stabilized, the resulting vegetation community is most dominated by *Stipa comata*, *Artemisia frigida*, and *Selaginella densa*. The difference between stabilized blow-outs and stabilized dunes is seen in the occurrence of low lying shrubs on the former which are less common on stabilized dunes. The dominant species in dune depressions, which occur on stabilized dunes, are *Stipa comata*, *Calamovilfa longifolia* and *Agropyron spp.* [17]. Many population patches of hairy prairie-clover have been found in grassland and prostrate shrub stands within the study site, most commonly on low angled south and west facing slopes with partial vegetation cover and exposed sand [18].

This PFRA pasture covers an area of $1.23 \times 10^8 \text{ m}^2$ with a maximum altitude of 520 m. The Dundurn sand hills are characterized by annual temperature extremes with a mean annual temperature of 2.2°C and a mean of 14.8°C during the growing season from May to September [19]. The growing season is short with about 105 frost free days, high wind speeds (mean annual wind speed of 15.7 km hr^{-1}), high evaporation rates (mean annual evaporation rate of 650 mm), and low precipitation rates (mean annual precipitation rate of 350 mm) [19]. Prevailing winds in southern Saskatchewan are from the northwest resulting in a general dune orientation of northwest to southeast.

2.2. Data Set

2.2.1. Land Cover Classification

A SPOT5 panchromatic image (wavelengths 490 - 690 nm) at 2.5 m spatial resolution was acquired on April 13, 2007. This imagery was selected for its high spatial resolution. Three SPOT5 multispectral images (green,

red, Near Infra-Red (NIR), and Mid Infra-Red (MIR) bands) at 10 m spatial resolution were acquired on May 27, 2009; July 12, 2009; and August 16, 2007. These images were selected for the multiple bands and covering three stages of the growing season.

A multi-temporal, multi-resolution land cover classification was carried out using object-oriented methods. Class boundary delineation was completed using a multi-resolution segmentation algorithm at a scale parameter of ten, with shape and compactness weighted at 0.1 and 0.5 respectively. Following segmentation, the landscape was classified in to the five main land cover types present in the study site: aspen, shrub, juniper, grassland, and bare sand with a sixth class, other, added during post classification editing. A nearest neighbour classifier was set up using twenty six training data points collected in the field for each class and class membership values were assigned based on the mean reflectance within the panchromatic, red, NIR, and MIR bands of an object. A hierarchical rule set was developed to maximize spectral and shape separability between classes.

The raster data format of the land cover classification was converted to vector data, retaining the original boundaries of raster data. Bare sand patches were extracted from the classification and dissolved to create a separate layer of the class bare sand. The remote sensing data set consisted of the bare sand layer.

2.2.2. Occupancy Data for Hairy Prairie-Clover

Occupancy data for hairy prairie-clover was collected between 2007 and 2010. A previous coarse resolution land cover map from Environment Canada was used to identify 351 patches within which to restrict ground-based searches for hairy prairie-clover. Systematic searches of these patches were conducted from July to September each year to allow for easy detection as the plant is in bloom during this time period. Searches included a complete search of the bare sand patch and a buffer area of 30 m surrounding each patch. If plants were found, the outer boundary of the population was tracked using GPS ($\pm 5 \text{ m}$ error) to record the spatial extent of the patch. In spatial analyses, pseudo replication of a variable will depend on the distance between the sample points, such that a set of closely spaced sample points will be spatially autocorrelated and thus decrease the effective sample size [2]. Therefore, to minimize spatial autocorrelation issues, any hairy prairie-clover plants that were found within a 30 m proximity of one another were counted as the same patch/population.

This polygonal occupancy data was overlaid on the bare sand layer extracted from the above land cover classification to create an attribute table for bare sand patches where hairy prairie-clover was present or absent. Presence/absence data from 2007-2009 was used as training

data in model building while presence/absence data collected in 2010 was kept separate for use as validation data in accuracy assessment. Reference [21] found that regular and equal-stratified sampling designs were the most accurate and robust for habitat suitability and presence/absence predictions. However, gradients that exert major control over a species distribution should be used to stratify sampling and increase efficiency [21]. Therefore, in this case it was most appropriate to use systematic searches of targeted bare sand areas in data collection.

2.3. Metrics for Bare Sand Patch Configuration

To determine inter and intra patch spatial characteristics of bare sand patches, patch scaled metrics were calculated. Metrics were chosen based on the premise that the likelihood of a species occupying a particular habitat niche is a function of patch area and patch isolation [6]. In plant species, dispersal can be affected by patch isolation, establishment by the availability of suitable habitat and habitat quality, and persistence by patch size [8]. Therefore, six metrics were calculated for bare sand patches: area, shape, fractal dimension, Euclidean nearest neighbour distance, proximity count, and proximity percent. The normalized difference vegetation index (NDVI) was also calculated as a measure of habitat quality.

Shape was calculated as

$$SHAPE = \frac{0.25 p_{ij}}{\sqrt{a_{ij}}} \quad (1)$$

where p_{ij} is the patch perimeter and a_{ij} is the patch area, adjusted by the constant 0.25 to adjust for the square standard in raster data [22]. Shape will equal one if a patch is perfectly square and increase without limit as patch shape becomes more irregular [22]. It is a unitless measure. Fractal dimension, another measure of patch shape, was calculated as

$$FRACT = \frac{2 \ln(0.25 p_{ij})}{\ln a_{ij}} \quad (2)$$

where p_{ij} is the patch perimeter, a_{ij} is the patch area, 0.25 is the constant for the square standard in raster data, and \ln is the natural logarithm [22]. Fractal dimension will vary between a value of one and two (unitless), where one signifies a square patch and two is the upper limit of shape irregularity [22].

The Euclidean nearest neighbour (ENN) distance was calculated as the straight line distance (meters) between the centroids of the closest neighbouring patches. Bare sand patch centroids were created based on the gravimetric center of patches. Proximity measures, as a representation of patch isolation, were based on a 30 m buffer

zone surrounding bare sand patches. Proximity count is the number of bare sand patches that fall within the 30 m buffer zone of the target sand patch. Proximity percent is the percentage of the area of the 30 m buffer zone that is of the bare sand class. These measures were modified from the proximity index by [22] to create a more meaningful output. Given that hairy prairie-clover spreads mainly by seed [18], it is plausible that dispersal distances fall within this proximity.

In plant species, one aspect of habitat quality is the competition for resources with other flora species, especially in inferior competitors because plants are immobile [3]. Therefore, it is important to consider the amount of vegetation growing on bare sand patches as one measure of habitat quality for hairy prairie-clover. NDVI was used as a surrogate measure of the amount of vegetation on a bare sand patch. NDVI can be calculated for an image on a per-pixel base as

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (3)$$

The difference in reflectance between the red band (absorption in vegetation, reflectance in soil) and NIR band (reflectance in vegetation and soil) will enhance vegetation signal across the imagery. NDVI is one of the most popular vegetation indices because it can capture seasonal changes in vegetation growth, it is highly correlated with leaf area index, and it works best in low biomass conditions such as grassland, semi-arid, and arid regions [23]. NDVI was calculated for the SPOT5 multispectral images from August and May. NDVI from August was chosen because during the late growing season, the maximum amount of vegetation coverage on a bare sand patch will show up [23]. NDVI from May was also chosen to show the amount of vegetation on a bare sand patch in the early growing season. Competition for resources in the early growing season can affect seedling establishment and persistence which is important for population growth [3,24]. The average NDVI for each sand patch was calculated.

2.4. Statistical Analysis and Model Building

Bare sand patches were coded based on hairy prairie-clover occupancy data where zero was unoccupied, one was occupied, and two was unknown. The data set consisted of 314 occupied and 314 unoccupied bare sand patches randomly selected from the 2007-2009 occurrence data. Using a binary response variable in logistic regression is a popular approach to habitat suitability modelling [21]. However, sampling of the data should be done in a way to reduce model bias. Therefore, equal sample sizes were chosen to represent occupied and unoccupied sand patches. The independent variables tested

were area, shape, fractal dimension, Euclidean nearest neighbour distance, proximity count, proximity percent, mean NDVI August, and mean NDVI May. For all tests, $\alpha = 0.05$ was used. Assumptions of normality were tested using the one-sample Kolmogorov-Smirnov test. For independent variables in normal distribution, the independent-samples t-test was used to test for a significant difference between occupied and unoccupied bare sand patches. Assumptions of equal variances were tested using Levene's test for equality of variances. The independent-samples Mann-Whitney U test was used to test for a significant difference between occupied and unoccupied bare sand patches for variables not in normal distribution [25].

Binary logistic regression was used to determine which independent variables were significantly correlated with hairy prairie-clover occurrences, because the dependent variable was binary and the independent variables were continuous [25]. Logistic regression is advantageous because it provides an equation for predicting the probability of occupancy on a bare sand patch, it uses multivariate comparisons between occupied and unoccupied sand patches, and it can identify the variables most associated with the occurrence of hairy prairie-clover [26]. The stepwise selection of independent variables can be useful for selecting an optimal model because the full model may be oversized, overfitted, or redundant [2]. Therefore, a forward stepwise likelihood ratio method was used (enter 0.05, removal 0.10). Correlation between independent variables can influence the final model because the selection and contribution of a variable depends on other variables already in the model [26]. Therefore, to avoid multicollinearity, only shape was used in the logistic regression and not fractal dimension. Mean NDVI August was also left out of the model because there was no significant difference between occupied and unoccupied bare sand patches for this variable ($P > 0.05$). Variables considered in the model were area, shape, Euclidean nearest neighbour distance, proximity count, proximity percent, mean NDVI May, and all possible interactions. The Hosmer-Lemeshow test was used to test the assumption that there is a linear relationship between the independent variables and the log odds of the dependent variable [25].

2.5. Habitat Suitability Predictions and Model Validation

The logistic regression equation was subsequently used to calculate the probability of occupancy for all bare sand patches. Probability of occupancy was calculated as

$$P = \frac{\text{Odds}}{(1 + \text{Odds})} \quad (4)$$

where odds was defined as

$$\text{Odds} = e^{a+bx} \quad (5)$$

with $a + bx$ being the logistic regression equation. Hairy prairie-clover occupancy data from 2010 was used for accuracy assessment of model predictions. The data set consisted of 377 unoccupied and 131 occupied sand patches. Hairy prairie-clover occurrence polygons were overlaid with the bare sand layer to create a contingency table of the model predictions (predicted suitable/unsuitable) against actual observations (observed presence/absence) (**Table 1**).

The Kmax index is the highest kappa value achieved when calculating the kappa coefficient at all threshold values between zero and one [1,27]. It is a threshold-independent evaluator. Contingency tables were constructed for threshold values 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. Threshold values are defined in **Table 2**. The kappa coefficient was calculated for each threshold value as

$$KIA = \frac{N \sum x_{ii} - \sum x_{i.} x_{.i}}{N^2 - \sum x_{i.} x_{.i}} \quad (6)$$

where x and N are counts of evaluation points as defined in **Table 1** [27]. The percentage of presences correctly predicted was calculated from the contingency tables as

$$\text{sensitivity} = (x_{ii}/x_{.i}) \times 100 \quad (7)$$

Table 1. Contingency table for calculation of KIA, sensitivity, specificity, and overall %.

| | | Predicted | | |
|----------|--------------|----------------|--------------|----------|
| | | Unsuitable (0) | Suitable (1) | Sum |
| Observed | Absence (0) | x_{ss} | x_{is} | $x_{.s}$ |
| | Presence (1) | x_{si} | x_{ii} | $x_{.i}$ |
| | Sum | $x_{.s}$ | $x_{.i}$ | N |

Table 2. Predicted suitable and unsuitable hairy prairie-clover habitat based on probability of occupancy for threshold values varying between zero and one. P represents the probability of occupancy for a bare sand patch and varies between zero and one.

| Threshold Value | Suitable Habitat | Unsuitable Habitat |
|-----------------|------------------|--------------------|
| 0.1 | $P \geq 0.1$ | $P < 0.1$ |
| 0.2 | $P \geq 0.2$ | $P < 0.2$ |
| 0.3 | $P \geq 0.3$ | $P < 0.3$ |
| 0.4 | $P \geq 0.4$ | $P < 0.4$ |
| 0.5 | $P \geq 0.5$ | $P < 0.5$ |
| 0.6 | $P \geq 0.6$ | $P < 0.6$ |
| 0.7 | $P \geq 0.7$ | $P < 0.7$ |
| 0.8 | $P \geq 0.8$ | $P < 0.8$ |
| 0.9 | $P \geq 0.9$ | $P < 0.9$ |

where x_{ii} and $x_{.i}$ are defined in **Table 1**. The percentage of absences correctly predicted was calculated from the contingency tables as

$$\text{specificity} = (x_{ss}/x_{.s}) \times 100 \quad (8)$$

where x_{ss} and $x_{.s}$ are defined in **Table 1**. The overall percentage correctly predicted was calculated from the contingency tables as

$$\text{Overall \%} = [(x_{ss} + x_{ii})/N] \times 100 \quad (9)$$

where N is defined in **Table 1**.

The area under curve (AUC) is one of the most common threshold-independent methods for evaluating species' distribution models because it allows one to weight the risks of over and under predicting species' presences [1,27]. For each threshold value the true-positive ratio was plotted against the false-positive ratio and the area under the resulting curve was calculated. As well, a second curve was plotted with the true-negative ratio against the false-negative ratio for calculation of the area under the curve. The true-positive ratio was calculated from the contingency tables as

$$\text{true positive} = x_{ii}/x_{.i} \quad (10)$$

as defined in **Table 1** [27]. The false-positive ratio was calculated from the contingency tables as

$$\text{false positive} = x_{is}/x_{.s} \quad (11)$$

as defined in **Table 1** [27]. The true-negative ratio was calculated from the contingency tables as

$$\text{true negative} = x_{ss}/x_{.s} \quad (12)$$

as defined in **Table 1** [27]. The false-negative ratio was calculated from the contingency tables as

$$\text{false negative} = x_{si}/x_{.i} \quad (13)$$

as defined in **Table 1** [27]. The AUC value will vary between zero (worse than random model) and one (best possible model), with values of 0.70 and greater indicating average to excellent model accuracy [2].

The predicted-to-expected (P/E) ratio was the final evaluation method used to assess model accuracy because it is threshold-independent, it is an assessment of presence-only data, and it can be used to predict several levels of habitat suitability other than just suitable and unsuitable [27]. Ten habitat suitability classes were defined based on the threshold values between zero and one (**Table 3**). For the P/E curve, ten classes with a class width of 0.1 is the optimum [27]. For each habitat suitability class, the P/E ratio was calculated as

$$Fi = Pi/Ei \quad (14)$$

where Pi is the predicted frequency and Ei is the expected frequency [27]. The predicted frequency was calculated as

Table 3. Habitat suitability classes defined in the P/E curve. P represents the probability of occupancy for a bare sand patch and varies between zero and one.

| Habitat Suitability Class | Probability of Occupancy |
|---------------------------|--------------------------|
| 0.0 | $P = 0.0$ |
| 0.1 | $0.0 < P \leq 0.1$ |
| 0.2 | $0.1 < P \leq 0.2$ |
| 0.3 | $0.2 < P \leq 0.3$ |
| 0.4 | $0.3 < P \leq 0.4$ |
| 0.5 | $0.4 < P \leq 0.5$ |
| 0.6 | $0.5 < P \leq 0.6$ |
| 0.7 | $0.6 < P \leq 0.7$ |
| 0.8 | $0.7 < P \leq 0.8$ |
| 0.9 | $0.8 < P \leq 0.9$ |
| 1.0 | $0.9 < P \leq 1.0$ |

$$Pi = \frac{pi}{\sum_{j=1}^b pj} \quad (15)$$

where pi is the number of validation data points that fall within the habitat suitability class i and $\sum pj$ is the total number of validation data points [27]. The expected frequency was calculated as

$$Ei = \frac{ai}{\sum_{j=1}^b aj} \quad (16)$$

where ai is the area covered by class i and $\sum aj$ is the total area of the study site [27]. In this case, the study site refers to bare sand patches because only habitat is being considered and not total landscape. Fi was plotted against habitat suitability to produce the P/E curve. It is expected that a good model will show a linearly increasing curve with a low suitability class containing fewer presences than expected by chance ($Fi < 1$) and high suitability classes having Fi increasingly higher than one. A threshold occurs where $Fi = 1$ below which habitat is unsuitable and above which habitat is suitable [27].

Because the P/E curve did not match what was theoretically expected, k-fold cross-validation was used to assess any biases that may be caused by how the data was used. Hairy prairie-clover presence/absence data was randomly divided into three independent partitions ($k = 3$) [27]. K-fold cross-validation was applied using two of the partitions to calibrate the model and the left out partition in accuracy assessment, on only the retained variables from the original model [27]. Therefore, two additional evaluations of model accuracy were produced to compliment the original accuracy assessment. Central tendency and variance of the P/E curve were evaluated from the mean, 95% confidence interval, and standard

deviation to assess model robustness [27]. The 95% confidence interval was calculated as

$$95\% CI = \bar{X} \pm t_{0.05}(2) \times S\bar{X} \quad (17)$$

where \bar{X} is the mean and $S\bar{X}$ is the standard error [25]. Standard error was defined as

$$S\bar{X} = \frac{S}{\sqrt{n}} \quad (18)$$

where S is the standard deviation and n is the sample size [25]. The standard deviation was calculated as

$$S = \sqrt{\frac{SS}{n-1}} \quad (19)$$

where the sum of squares (SS) was defined as

$$SS = \sum (X_i - \bar{X})^2 \quad (20)$$

The width of the confidence interval can assess the model's sensitivity to particular calibration points [27].

Bare sand patches were then re-coded as their predicted habitat suitability based on the probability of occupancy. The habitat suitability threshold for dividing suitable habitat (probability of occupancy above threshold) from unsuitable habitat (probability of occupancy below threshold) was chosen as the threshold value with the highest kappa coefficient, and where the true-positive ratio was maximized and the false-negative ratio was minimized [27]. When choosing the habitat suitability threshold, the goal was to minimize the false-negative ratio so as not to miss valuable hairy prairie-clover habitat. Additionally, four habitat suitability classes were defined from the steps in the P/E curve [27].

3. RESULTS AND DISCUSSION

3.1. Inter-Patch and Intra-Patch Spatial Characteristics for Potential Bare Sand Habitat

Independent-samples t-tests and Mann-Whitney U tests revealed that there was a significant difference between occupied and unoccupied sand patches for most of the variables selected (**Table 4**). Occupied sand patches were significantly larger in size, more elongated shaped, less isolated, and less vegetated in the early growing season ($P < 0.05$). There was no significant difference between the NDVI of occupied and unoccupied sand patches in the late growing season ($P > 0.05$). Therefore, there is no evidence that late season vegetation cover on sand patches significantly affects the occurrence of hairy prairie-clover. Although, there is evidence that vegetation cover on sand patches in the early growing season affects the occurrence of hairy prairie-clover. The mean NDVI in May was significantly lower on occupied sand patches as opposed to unoccupied sand patches (**Table 4**). Seed-

ling establishment and persistence in the early growing season can be affected by competition for resources with other species, therefore it may be advantageous to hairy prairie-clover to establish on bare sand areas with reduced vegetation cover in the early growing season [3]. As well, different types of vegetation green-up at different times during the growing season, therefore the NDVI in May could be detecting differences in the type of vegetation on sand patches rather than differences in the amount of vegetation. For example, it is commonly observed in the field that juniper inhabits partially to heavily stabilized bare sand areas. Considering juniper is an evergreen shrub it will be captured in the early season NDVI as opposed to species that have not undergone green-up at this time. The difference in NDVI may be capturing the difference in juniper occurrence between sites.

Unoccupied bare sand patches tended to be smaller, more round shaped, more isolated, and more vegetated in the early growing season. Unoccupied sand patches had closer neighbouring sand patches than occupied bare sand areas but, were surrounded by less neighbouring patches. The ENN distance may be a poor measure of isolation because larger patches will have a greater distance between their patch centroids compared to smaller patches. Thus, the value may be an artefact of the measurement and not a true measure of isolation. This also suggests that it is not only the distance between habitat patches but the amount of surrounding habitat that is important to hairy prairie-clover occurrence. Dispersal is mainly affected by patch isolation and matrix quality with small patches being more affected by the quality of the matrix than larger patches [7,8,28]. However, small patches that more densely populate the matrix can influence the nature of the intervening matrix to decrease patch isolation [28]. Therefore, the amount of bare sand habitat within a 30 m proximity of the target patch may be a more ecologically relevant measure of patch isolation than just the ENN distance when considering factors affecting dispersal.

Table 4. Significant differences between occupied and unoccupied sand patches for chosen patch metrics.

| Variables | Mean (Occupied) | Mean (Unoccupied) | Significance |
|------------------------|-----------------|-------------------|--------------|
| Area (m ²) | 724 | 188 | P < 0.05 |
| Shape | 1.69 | 1.35 | P < 0.05 |
| Fractal Dimension | 1.16 | 1.11 | P < 0.05 |
| ENN Distance (m) | 32 | 28 | P < 0.05 |
| Proximity (count) | 4 | 3 | P < 0.05 |
| Proximity (%) | 15.76 | 12.34 | P < 0.05 |
| NDVI May | 0.38 | 0.41 | P < 0.05 |
| NDVI August | 0.65 | 0.65 | P > 0.05 |

3.2. The Spatial Configuration of Hairy Prairie-Clover Habitat

Binary logistic regression revealed that the size and shape of a sand patch, as well as the degree of isolation, and amount of vegetation in the early growing season were significantly correlated with hairy prairie-clover occurrences (Eq.21). Almost 29% of the variation in bare sand patch occupancy could be explained by the logistic regression equation ($r^2 = 0.285$ and $P < 0.05$).

$$\ln(\text{ODDS}) = 4.069 - 0.013\text{Area} + 0.036\text{ProxPbyShape} - 13.582\text{NDVIMay} + 0.038\text{NDVIMaybyArea} \quad (21)$$

The variable interaction of proximity percent with shape was the most useful for differentiating between occupied and unoccupied sand patches (Table 5). Similarly, other studies have found that the isolation of habitat patches is most important in explaining species richness and community composition because the rescue-effect allows populations to persist in small patches if they are well connected [29]. Adding proximity percent, shape, area, and NDVI May significantly improved the model's ability to predict suitable and unsuitable bare sand habitat for hairy prairie-clover (Table 5).

Sand patch area and NDVI May could be significant because larger sand patches tend to be less vegetated which decreases the competition for resources with other plants. Biotic interactions between plants act at short distances, but plants are immobile and have limited dispersal, therefore it is advantageous for inferior competitors to occupy competition-free locations [3]. Although some vegetation does take hold on the slipface of large sand dunes, these areas are generally kept clear of vegetation due to wind erosion and lack of soil nutrients and moisture favourable to the growth of most plant species [15], thus providing favourable habitat and competition-free locations for hairy prairie-clover.

Considering that herbivores are one of the major vectors of seed dispersal within the study site [18], proximity percent and sand patch shape can help explain the spatial distribution of hairy prairie-clover. For example,

Table 5. Results of logistic regression analysis of variables characterizing occupied and unoccupied sand patches.

| Variable Entered | Model Chi-Square | Improvement Chi-Square |
|------------------|------------------|------------------------|
| ProxP by Shape | 113.56* | 23.82* |
| NDV IMay | 129.49* | 15.93* |
| NDVI May by Area | 139.94* | 10.45* |
| Area | 152.26* | 12.32* |

*Significant at $P < 0.01$.

ungulates can exhibit specific dispersal behaviour such as the orientation of movement along landscape elements and the ability to choose between patches, which will affect dispersal and exchange rates of seeds between patches [30]. Ungulates, such as white-tailed deer, are more likely to prefer linearly shaped patches because they provide more preferential edge habitat and corridors for movement [31,32]. Thus, the shape of a sand patch could indirectly affect the spatial distribution of hairy prairie-clover through foraging behavior within the landscape. It was found that occupied sand patches were significantly more elongated than unoccupied sand patches (Table 4).

As well, the percentage of the immediate surrounding area (30 m buffer) of a sand patch could indirectly influence the spatial distribution of hairy prairie-clover through its influence on seed dispersal. Dispersal and exchange rates of seeds between patches will flow more readily in well connected patches based on the behavioural rule where individuals will stay at the first patch they reach as opposed to travelling extended distances between patches [30]. Occupied sand patches were also significantly less isolated than unoccupied sand patches (Table 4). Therefore, the spatial distribution of hairy prairie-clover can be explained by three constraints: 1) Patch Area (Area): the local environment allows the population to grow, 2) Patch Quality (NDVI May): interactions with other local species allow the species to persist, and 3) Patch Isolation (ProxP and Shape): the location is accessible given the dispersal abilities of the species [3].

3.3. Habitat Suitability Predictions

The habitat suitability threshold was set at 0.40 because the kappa coefficient was maximized and the sensitivity of prediction was substantially higher than that achieved at a threshold of 0.50 (Table 6 and Figure 2). If the probability of a sand patch being occupied by hairy prairie-clover was equal to or greater than 40%, the sand patch was predicted to be suitable habitat. If the probability of occupancy was less than 40% then the sand patch was predicted to be unsuitable habitat. The goal was to minimize the false-negative ratio by maximizing the sensitivity of predictions. A desirable habitat suitability threshold occurred at 0.40 when considering the risks of over and under predicting suitable habitat (Figure 2). At a threshold value of 0.40, the true-positive ratio is maximized while the false-negative ratio is minimized. At this threshold, 35.5% of the sand patches in the study site were predicted to be suitable hairy prairie-clover habitat and 64.5% were predicted to be unsuitable habitat.

The results of the accuracy assessment for habitat suitability predictions show an average level of model

accuracy (Table 6 and Figure 2). The highest kappa value obtained for the logistic regression model was 0.25, indicating a fair level of model agreement with ground referenced data [33]. As well, AUC value at 0.70 indicates an average level of model accuracy [2]. Reference [1] found similar AUC values in their study comparing the accuracy of different species' distribution models. Across all species in all regions studied and considering only the best methods, they found that 14% of the species had AUC values around 0.70. At the chosen habitat suitability threshold, about 61% of the time the model will correctly predict suitable and unsuitable habitat; about 80% of the time, occurrences are correctly predicted; and about 54% of the time, absences are correctly predicted (Table 6).

3.4. Relationship between Habitat Pattern and Population Pattern

The P/E curve was used to determine threshold values for four habitat suitability classes (Figure 3). A low suitability class should contain fewer evaluation presences than expected by chance ($Fi < 1$) and high suitability classes should have Fi increasingly higher than 1 [27]. Similar to the results found with the Max Kappa and AUC methods, the habitat suitability threshold for separating suitable habitat from unsuitable habitat occurred at 0.40 where the P/E curve crosses the $Fi = 1$ line (Figure 3). Unsuitable habitat occurred across a probability of

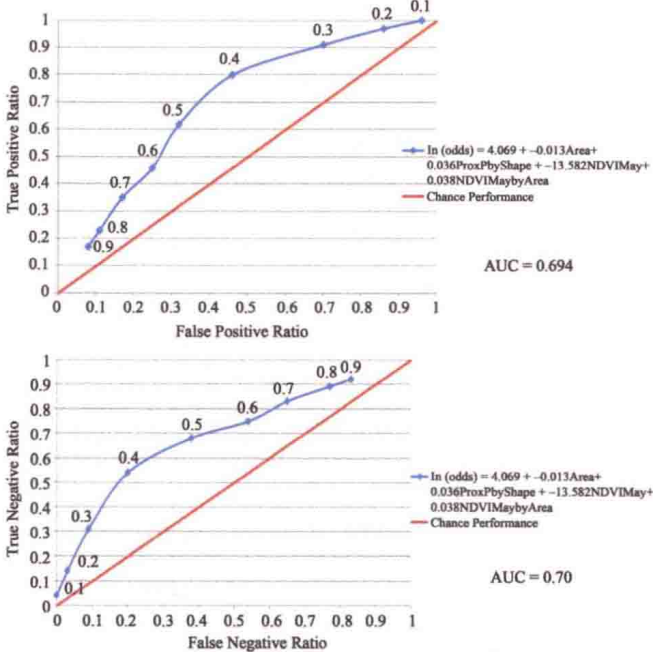


Figure 2. Accuracy assessment of habitat suitability predictions using the Area Under Curve (AUC) method. The red line represents a model that is no better than chance performance and the blue line represents the model derived from the logistic regression equation. The data points on the blue line represent the habitat suitability threshold values as defined in Table 2.

occupancy less than or equal to 20% where there were fewer evaluation presences than expected by chance (Figure 3). Marginally unsuitable habitat occurred near the transition between suitable and unsuitable habitat at the $Fi = 1$ line where the number of evaluation presences was close to what would be expected by chance (Figure 3). Suitable habitat occurred where the model predicted the probability of occupancy to be between 40% and 88%, because there were more evaluation presences than expected by chance (Figure 3). Where the model predicted the probability of occupancy to be greater than 88% (marginally suitable), validation data showed that occupancy of these types of sand patches actually dropped off which was not what was theoretically expected to occur (Figure 3).

Table 6. Accuracy assessment of habitat suitability predictions using the MaxKappa method. KIA is the kappa coefficient. Threshold values are based on the probability of occupancy as derived from Eq.21. Threshold values are defined in Table 2.

| Threshold | KIA | Sensitivity of Prediction | Specificity of Prediction | Overall % |
|-----------|------|---------------------------|---------------------------|-----------|
| 0.1 | 0.02 | 100 | 3.7 | 28.5 |
| 0.2 | 0.06 | 97.0 | 14.3 | 35.6 |
| 0.3 | 0.13 | 90.8 | 30.5 | 46.1 |
| 0.4 | 0.25 | 80.2 | 53.8 | 60.6 |
| 0.5 | 0.25 | 61.8 | 67.9 | 66.3 |
| 0.6 | 0.20 | 45.8 | 75.3 | 67.7 |
| 0.7 | 0.19 | 34.8 | 82.8 | 70.5 |
| 0.8 | 0.14 | 22.7 | 89.1 | 72.0 |
| 0.9 | 0.10 | 16.7 | 92.0 | 72.4 |

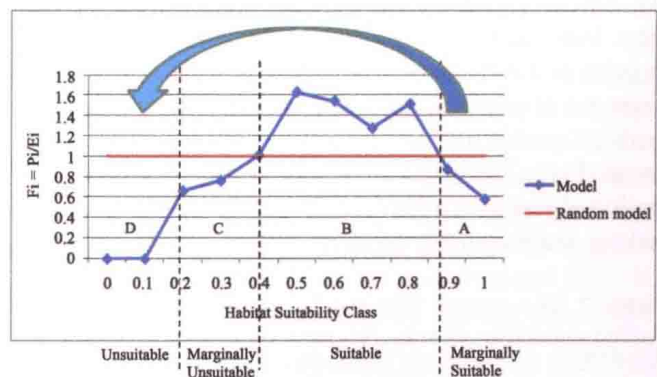


Figure 3. Four habitat suitability classes as derived from the P/E curve. The blue line represents the P/E curve derived from the logistic regression model and the red line represents a chance or random model. Fi is the predicted to expected ratio and the habitat suitability classes are defined in Table 3. The blue arrow represents the direction of natural succession when considering two variables from the logistic regression model: area and NDVI May. (A) Active sand areas (Marginally Suitable habitat: increasing false absences); (B) The transition state of a partially stabilized sand patch (Suitable habitat: mostly true presences); (C) The remnant state of a heavily vegetated sand patch (Marginally Unsuitable habitat: increasing false presences); (D) The altered state with the persistence of a new plant community (Unsuitable habitat: mostly true absences).

The shape of the P/E curve can be explained by how the logistic regression equation predicts the probability of occupancy. Looking at two variables, NDVI May and area, the model predicts that as sand patches become larger and NDVI May becomes lower the sand patch is more likely to be occupied by hairy prairie-clover (**Table 7**). Following the shape of the P/E curve, this is true until a threshold is reached around a probability of 0.88 where actual occupancy drops back to being below predicted occupancy. Considering NDVI May and area, natural succession on a sand patch would thus cause the probability of occupancy to decrease as colonizing species stabilize the bare sand area and gradually increase NDVI May and decrease area (**Figure 3**). This is apparent when looking at the average area and average NDVI May of each habitat suitability class (**Table 7**).

Landscapes are dynamic rather than static and are in a constant flux of change. The temporal scale of landscape change is different from the temporal scale of flora species' response to landscape change such that present species occurrences can still reflect past landscape patterns [29,34]. For example, [29] found that current species richness in fragmented alvar grasslands was more significantly related to past areas and connectivities of habitat patches than current habitat pattern. As well, [34] found a significant relationship between present habitat area and current species richness, and habitat area 50 years ago and current species richness in grasslands in south-eastern Sweden. Reference [35] found that remnant patches could exist within a population because there was a considerable time span between migration of new individuals to a patch and the final extinction of species in a patch during succession. Due to the delayed response of plants to habitat change, habitat patches can harbour species that prefer an environment that no longer exists [34]. Therefore, it is not uncommon that hairy prairie-clover was found to occupy marginally suitable habitat and marginally unsuitable habitat in this study.

Table 7. The average area and NDVI May of four different habitat suitability classes. As sand patch size increases and NDVI May decreases, the probability of occupancy increases.

| Predicted Habitat Suitability | Probability of Occupancy | Mean NDVI May | Mean Area (m ²) | % of Landscape |
|-------------------------------|--------------------------|---------------|-----------------------------|----------------|
| Unsuitable | $P \leq 0.2$ | 0.49 | 79.18 | 18.8 |
| Marginally Unsuitable | $0.2 < P \leq 0.4$ | 0.41 | 79.63 | 45.7 |
| Suitable | $0.4 < P \leq 0.88$ | 0.38 | 263.85 | 32.7 |
| Marginally Suitable | $P > 0.88$ | 0.41 | 2041.31 | 2.8 |

Considering the variables present in the logistic regression equation (**Table 7**) and the shape of the P/E curve, the four habitat suitability classes were derived based on the direction of natural succession (**Figure 3**). The sand dunes found in the Dundurn area comprise four physiographic categories: active complexes, stabilized blow-outs, stabilized dunes, and dune depressions. When considering three dominant environmental gradients, soil-water holding capacity, edaphic characteristics, and degree of stabilization; stabilized blow-outs and dune depressions have closely related environmental niches while stabilized dunes and active complexes are distinct and follow a distinct environmental gradient between the two. Natural succession reflects the stabilization of active complexes, increasing the water retaining capacity and organic content of sandy soils, towards stabilized dunes. Comparatively, blow-outs represent a more advanced stage of succession due to decreased wind erosion allowing invasion by woody species such as creeping juniper. As well, the more mesic conditions in dune depressions as compared to the exposed slopes of stabilized dunes, allows late successional species to take over and develop a more complex vegetative community [17].

On average, larger sized sand dunes are kept relatively clear of vegetation on their most active surfaces (the dune head, crest, and slip face) due to wind erosion and deposition of sand, and a lack of soil moisture and nutrients preventing growing conditions that are favourable to most species [15]. Active complexes in the Dundurn sand-hills were found to have lower soil organic content, higher percentages of sand, and lower soil moisture [17]. Although these types of sand patches were predicted to have a probability of occupancy greater than or equal to 88% for hairy prairie-clover, accuracy assessment revealed that *Fi* was less than one (**Figure 3(A)**). Therefore, these types of sand patches are only marginally suitable and the occurrence of false absences is increasingly higher in these types of sandy areas. False absences occur in these sandy areas because hairy prairie-clover tends to be absent from these potentially suitable sites (**Figure 3(A)**). While currently they are not preferred habitat, hairy prairie-clover plants are able to colonize them allowing succession to progress towards an environment that is more favourable to the plant.

Hairy prairie-clover prefers to occupy the niche in the successional stage where bare sand areas are partially stabilized (**Figure 3(A)** and **Table 7**). Thus, following the environmental gradient along the succession of active complexes to stabilized dunes [17], hairy prairie-clover will be more commonly associated with stabilized dunes than active complexes (**Figure 3**). Hairy prairie-clover has been found to be positively associated with early invading species of open sand and other early successional species within the Dundurn sand-hills [18]. Sparsely