Expansion Rate & Dispersal Pattern of the Non-Native Roesel’s Bush-Cricket in Sweden

Sonja Preuss

Faculty of Natural Resources and Agricultural Sciences
Department of Ecology
Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2012
Expansion rate & dispersal patterns of the non-native Roesel’s bush-cricket in Sweden

Abstract
Environmental change and anthropogenic activities influence species distributions. Species introductions have become increasingly common in an era of globalization and increased international trade and travel. The establishment of introduced species outside their native range and subsequent spread are of great conservation concern. Introduced species that become invasive, spread rapidly and reach high abundance, may cause the extinction of native species, disrupt ecosystem functioning and pose a threat to human health and the economy. It is therefore of great interest to understand the processes and mechanisms involved in species range expansion in order to develop effective management strategies. In this thesis I examine the influence of the landscape on species’ distribution and analyse patterns of range expansion of a non-native insect in south-central Sweden. Roesel’s bush-cricket (Metrioptera roeselii) was chosen as a model organism as its biology is well studied and its range expansion has been documented not only in Sweden but also in several other European countries. The aims of this thesis were (I) to identify landscape variables that predict the species distribution, (II) to estimate the rate of range expansion, (III) to identify the source of range expansion in south-central Sweden and to assess the dispersal pattern using population genetic data, and (IV) to analyse the influence of landscape composition and structure on population connectivity. I analysed species distribution, genetic and landscape data using a range of statistical modelling techniques in combination with geographic information systems (GIS). The results showed that the amounts of arable land, pasture and rural settlements as well as linear habitat elements are important predictors of the species’ distribution. During the last three decades, Metrioptera roeselii has expanded its range from the northern shores of the Lake Mälaren at an estimated rate of 0.3 - 3.16 km/year. The genetic diversity across the range was surprisingly high and degree of population differentiation was low to moderate likely due to frequent gene flow between populations in the centre of the species range and decreased gene flow towards the range margin. It appears the species establishes populations through infrequent long-distance and frequent short-distance dispersal (natural, human-mediated).

Keywords: distribution modelling, genetic diversity, gene flow, landscape analyses, Orthoptera, range expansion

Author’s address: Sonja Preuss, SLU, Department of Ecology, P.O. Box 7044, 750 07 Uppsala, Sweden
E-mail: Sonja.Preuss@slu.se
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Publications</td>
<td>7</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>11</td>
</tr>
<tr>
<td>2 Thesis Aim</td>
<td>15</td>
</tr>
<tr>
<td>3 Material and Methods</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Study area</td>
<td>17</td>
</tr>
<tr>
<td>3.2 The study species</td>
<td>17</td>
</tr>
<tr>
<td>3.3 Distribution of Roesel’s bush-cricket</td>
<td>20</td>
</tr>
<tr>
<td>3.4 Modelling methods</td>
<td>21</td>
</tr>
<tr>
<td>3.5 Population genetic structure</td>
<td>22</td>
</tr>
<tr>
<td>4 Results and Discussion</td>
<td>23</td>
</tr>
<tr>
<td>4.1 Species distribution modelling</td>
<td>23</td>
</tr>
<tr>
<td>4.2 Rate of range expansion</td>
<td>24</td>
</tr>
<tr>
<td>4.3 Local origin and genetic patterns across the range</td>
<td>27</td>
</tr>
<tr>
<td>4.4 Functional connectivity-dispersal routes</td>
<td>31</td>
</tr>
<tr>
<td>5 Conclusions</td>
<td>35</td>
</tr>
<tr>
<td>References</td>
<td>37</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>43</td>
</tr>
<tr>
<td>Tack!</td>
<td>45</td>
</tr>
</tbody>
</table>
List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:


III Sonja Preuss, Åsa Berggren, Anna Cassel-Lundhagen. Genetic variation and differentiation in a bush-cricket species under range expansion. (manuscript).

IV Åsa Berggren, Sonja Preuss, Göran Nordlander, Anna Cassel-Lundhagen. Examining the expansion pattern of Roesel’s bush-cricket by use of least-cost path analysis. (manuscript).

Paper I is reproduced with the permission of the publishers.
The contribution of Sonja Preuss to the papers included in this thesis was as follows:

I Main author, field work, analysis. Design and idea together with Åsa Berggren and Anna Cassel-Lundhagen.

II Main author and fieldwork.

III Main author, field work, laboratory work, and analyses. Design and idea together with Anna Cassel-Lundhagen and Åsa Berggren.

IV Co-author, field work and laboratory work. Design and idea together with Åsa Berggren and Anna Cassel-Lundhagen.
Abbreviations

BAPS  Baysian Analysis Population Structure
CMX  Conditional Maximum Model
Fst  Fixation index
GAM  Gamma Quantile Model
GC  Grid Count Model
GIS  Geographic Information System
GPS  Global Positioning System
GSD  Geographic Sweden Data
LCP  Least-Cost Path
M  Mean Model
MdN  Median Model
MR  Margin Model
MX  Maximum Model
1 Introduction

Understanding which factors and processes determine species distributions has been of prominent interest throughout the history of ecological research (Andrewartha and Birch 1954; Brown et al. 1996; Lockwood et al. 2007). With environmental change accelerating in recent decades, attention has been drawn to quantify the ongoing changes in our ecosystems and to understand the driving forces (Mack et al. 2000; Sala et al. 2000; Parmesan and Yohe 2003; Brooker et al. 2007). There is mounting evidence that species’ distributions have changed in response to climate change and anthropogenic activities (Parmesan et al. 1999; Jetz et al. 2007; Kornis et al. 2012; Zoazaya et al. 2012). Globalization and increased international trade have contributed to the introduction of species outside their historic ranges by transporting them across major geographical dispersal barriers (Wilson et al. 2009). Species introductions are often associated with undesired ecological and economic impacts. Biological invaders are now widely recognized as one of our most pressing conservation threats (Vitousek et al. 1996; Mack et al. 2000). The uncontrolled spread of introduced species can cause habitat degradation, extinction of native flora and fauna, changes in ecosystem functioning, and facilitation of subsequent invasions. Invasive species threaten 60% of red-listed species in US (Arim et al. 2006). Large economic impacts are also associated with many invasive species, which can provoke agricultural losses, disrupt ecosystem services, and lead to disease proliferation (Andow et al. 1990; Pimentel et al. 2005). Research efforts have therefore been directed to understand and predict invasion dynamics and dispersal patterns; knowledge which is essential for the development of effective conservation and management strategies to limit and control and the spread of undesired species (Hulme 2003).

Biological invasions are commonly associated with negative connotations amongst the wider audience, based on the known or believed environmental or
socio-economic impacts that can be caused by invasive species (Valery et al. 2008). However, impacts of introduced species are not always proven to be negative and not all non-native species *per se* become invasive (e.g. Ricciardi and Cohen 2007). It is a matter of debate when introduced species are considered naturalized (Richardson et al. 2000), and there are even examples of native species becoming ‘invasive’, i.e. becoming both widespread and locally dominant (Richardson et al. 2011). It is important to note that ‘invasiveness’ is not a fixed species attribute as species dispersal patterns are dynamic and strongly influenced by environmental conditions (Colautti and MacIsaac 2004).

In the context of the present study (thesis) I adopt a neutral, process-based definition of ‘invasive’ which refers to the organism’s potential to rapidly colonize a large area, and to focus on aspects of dispersal behaviour, rate of spread and establishment success. This neutral definition of ‘invasive’ has been previously advocated (Richardson et al. 2000; Colautti and MacIsaac 2004). The argument is that invasions represent biogeographical rather than taxonomic phenomena and invasion success is determinate at the population level and not at the species level (Thebaud and Simberloff 2001; Torchin et al. 2002).

Invasions can be described as a multi-stage process that can be broadly divided into three phases: I) colonization II) establishment and III) expansion. In the colonization phase the organism needs to overcome a geographic barrier in order to colonize distant habitats. At this stage the persistence of the species depends on repeated introductions, with the colonization success being governed by dispersal opportunities and propagule size (Kolar and Lodge 2001). In the establishment phase, a reproductive barrier is overcome and species persistence is no longer dependent on repeated introductions. Species with a wide niche breadth, that are able to adapt changed environmental conditions and can rapidly achieve high densities, may have greater establishment success (Kolar and Lodge 2001). During the expansion phase, the species spreads into the surroundings with increasing numbers of individuals leaving the natal population to colonize distant habitat patches. Species which are highly dispersive and can acquire a competitive advantage are likely to expand rapidly and become locally dominant.

The invasion process is additionally influenced by environmental conditions and population dynamics. The rate of range expansion can be weakened by biological constraints such as reduced genetic diversity in small founder populations. Accidental introductions of species and the onset of range expansions are difficult to detect as they often occur over large temporal and spatial scales. The time lag between the colonization and invasion phase makes it difficult to detect species’ range expansions at an early stage (Mack et al.
2000; Ricciardi and Cohen 2007). This applies particularly to small and inconspicuous species that disperse via small propagules (Hulme 2009). In this thesis I use Roesel’s bush-cricket (Metrioptera roeselii) as a model organism to study the dispersal and range expansion patterns of a non-native species in south-central Sweden.
2 Thesis Aim

The current range expansion of Roesel’s bush cricket (*Metrioptera roeselii*) in south-central Sweden was chosen as a model system to improve our understanding of what determines the distribution and dispersal of an expanding species. Such knowledge is relevant for the management of introduced species, survey design and development of management strategies for native and non-native species.

The main objectives were:

- To identify landscape variables that predict the species’ large scale distribution (Paper I)

- To estimate the rate of range expansion and compare the performance of seven established methods for measuring range expansion regarding their sensitivity to sampling effort (Paper II)

- To identify the species’ local origin and analyse the spatial patterns of range expansion and recurrent dispersal (Paper III & IV)
3 Material and Methods

3.1 Study area

The landscape of south-central Sweden consists of a mosaic of agricultural land (46%), forest (43%), scattered settlements and small towns (5%), lakes and waterways (3%) and a small proportion of marshes and rocky outcrops (3%). The study area is situated in the Lake Mälaren region (mid-point 59°35’N, 16°30’E) between the inland harbour cities of Västerås and Örebro and extends over an area of approximately 120 x 140 km (Figure 1).

3.2 The study species

Roesel’s bush-cricket is a suitable study species for the following reasons: *Metrioptera roeselii* is non-native to Sweden and has expanded in its range during the last decades, not only in Sweden but also in other European

*Metrioptera roeselii* (Orthoptera: Tettigoniidae) (Hagenbach 1822) is a small (12-18 mm) omnivorous generalist species commonly found in tall grassland habitats across central and eastern Europe (Bellmann 2006; Eades et al. 2012). In Sweden the species was first sighted in 1882 near the inland harbour city Västerås at the shore of Lake Mälaren (Brunner von Wattenwyl 1882). Historical trade records and recent genetic analysis suggest that the species was introduced from northern Germany during the 19th century (Bäckström 1924; Montelius 1993; Kaňuch et al. unpublished). For nearly one hundred years the species went largely unnoticed. First observations on the local spread of the species were made in 1970 (Idar 1973), followed by comprehensive surveys in 1989 - 1990 and 2008 - 2010 which documented the range expansion of *M. roeselii* in the region (de Jong and Kindvall 1991; Species Gateway 2010; Preuss et al. 2011). Over the last 20 years the species has continued to spread rapidly at an estimated rate of 1 - 3 km/year, resulting in an observed shift of the range margin by 50 km (Paper II).

![Figure 2. *Metrioptera roeselii*, male, short-winged. Photo: Sonja Preuss](image)

*Metrioptera roeselii* is a wing dimorphic species with complex dispersal behaviour. The majority of the individuals (98 - 99 %) are usually short-winged (Figure 2) and disperse short distances by walking and jumping (Vickery 1965; Wissmann et al. 2009). The observed movement distances of
short-winged individuals ranges from 7 to 172 m/day (Kindvall et al. 1998). The species avoids heavily grazed pastures and individuals move faster through grasslands with short vegetation (Berggren 2004; Gardiner and Haines 2008). The presence of conspecific males increases male movement rates probably to minimize competition over females (Berggren 2005). Favourable weather conditions (mild springs, hot summers) and high population densities, however, trigger the development of long-winged individuals (Figure 3) that are capable to fly and colonize distant habitat patches (Simmons and Thomas 2004; Gardiner 2009). The proportion of long-winged individuals can temporarily reach 20% but usually drops quickly within only few generations (Simmons and Thomas 2004; Poniatowski and Fartmann 2008), presumably due to the trade-off between increased mobility and reduced reproductive capacity in long-winged individuals (Zera and Denno 1997). The observed maximum value of the mean activity radius of long-winged individuals is 144.4 m (Poniatowski and Fartmann 2011), and in the year following long-distance dispersal average displacement distances of 10.8 km were observed (Hochkirch and Damerau 2009). Besides dispersing actively over short and long distances, passive dispersal by humans is likely to contribute to the range expansion of the species (Kaňuch et al. unpublished). *Metrioptera roeselii* colonizes different grassland habitats in the agricultural landscape and overwinters as eggs in hollow grass stems, which makes the species like other plant material ovipositing insects susceptible to human mediated translocation (Wagner 2004).

*Figure 3. Metrioptera roeselii*, male, long-winged. Photo: Sonja Preuss

In Sweden the species has univoltine or semivoltine life cycles, depending on the weather conditions and the time of egg laying (Ingrisch 1986; de Jong and Kindvall 1991). The nymphs hatch in May and go through six instars
before reaching maturity in July. Nymphs feed on flower parts and pollen while adults feed mainly on grass, grass seeds and small insects (Marshall and Haes 1988). Mature males stridulate on warm and sunny summer days nearly continuously to attract females (≥18°C; July to mid September). The characteristic song makes the species easy to census. *Metrioptera roeselii* prefers moist tall-grass areas but can also be found in other grassland habitats, such as extensively used meadows and pastures, fallow land, grassy field margins, ditches and road verges (Marshall and Haes 1988; Berggren et al. 2001; Poniatowski and Fartmann 2005; Gardiner and Hill 2008; Preuss et al. 2011). Forests, urban areas, arable crop fields and intensively grazed pastures are considered unsuitable habitat for the species (de Jong and Kindvall 1991; Ingrisch and Köhler 1998).

### 3.3 Distribution of Roesel’s bush-cricket

The Swedish distribution of Roesel’s bush-cricket in south-central Sweden was analysed based on records of the Swedish National Database (Species Gateway), available from 1981 onwards, and observational data from field surveys that were conducted between 1981 - 1990 and 2008 - 2010.

The Species Gateway (www.artportalen.se) is a publically accessible database which contains detailed data on species distributions, including information on geographical position, abundance and in some cases data on the life-history stage and biotope type where the observations were made. The data is collected and entered by volunteers, conservationists and researchers, and reports are checked by taxonomic specialists. The database contains 510 *M. roeselii* records on 2321 individuals from 455 locations that were reported in 17 years from 1981 - 2010.

For the field surveys known locations of *M. roeselii* were used as starting points to map the presence and absence of the species (within 5 x 5km grid cells) in the Lake Mälaren region (de Jong and Kindvall 1991; Berggren et al. 2001; Species Gateway 2010). Since the call of stridulating males is strong and can be heard over distances of > 10 m (Fischer et al. 1997; Bellmann 2006), it is possible to conduct auditory surveys by car, listening from the car window while driving slowly along countryside roads (~30 km/h; de Jong and Kindvall 1991; Berggren et al. 2001). The surveys were carried out on sunny days, between 10 am - 5 pm, from mid-July until the end of August. Survey routes and observations of *M. roeselii* were either noted on maps (1989 - 1990) or by using a handheld GPS (2008 - 2010) and later entered on digital maps in ArcGIS 9.2 (ESRI 2006).
The distribution data were used in the four studies presented in this thesis. Presence and absence data collected in 2008 and 2009 were used to model the species distribution with 1 x 1 km resolution (Paper I). All available presence data (1981 - 2010) and a resolution of 5 x 5 km was used to estimate the rate of range expansion of *M. roeselii* in south-central Sweden (Paper II). Based on large-scale distribution information (Species Gateway 2010), in 2010 we located the marginal populations and collected individuals for the population genetic analyses (Paper III & IV).

### 3.4 Modelling methods

I used binary logistic regressions to model the link between landscape variables and species presence - absence data to identify the most important environmental factors explaining the species distribution based on the generated statistical relationships (Paper I). The species - environment relationship was investigated at two spatial scales, in 1 x 1 km grid cells and within 10 m buffer area along surveyed roads. This was done to test for the impact of scale on model performance, which is important to consider when predicting current and future distributions of species (Cote and Reynolds 2002; Scott et al. 2002).

ArcGIS 9.2 (ESRI 2006) was used to plot the species distribution data and to extract relevant information on landscape composition from digital maps available from the Swedish mapping, cadastral and land registration authority (Geographic Sweden Data (GSD) topographic & land cover map 1:50 000, property map 1:10 000, and road map 1:100 000).

We used a least-cost modelling approach to assess the impact of landscape composition on species dispersal and to measure the effective distances among populations in the core area of the species range (Paper IV). Based on findings from study (I), and knowledge on species movement behaviour (Berggren et al. 2002; Berggren 2004) different hypothetical ‘costs’ were assigned to land use types depending on if they impede or facilitate the species movement. The paths that minimize the cumulative costs between populations, the ‘least-cost paths’, were identified using the Spatial Analyst package in ArcGIS. The influence of landscape composition on gene flow among populations was explored in isolation-by-distance analyses (Mantel tests) testing the correlation between population differentiation and least-cost path distances.

To estimate the rate of range expansion we tested and compared seven different models, one area based method and six distance based methods, using gridded occupancy data at a resolution of 5 x 5 km (Paper II). The rate of range
expansion was derived from the slope of the regression lines (Van den Bosch et al. 1990; Hassall and Thompson 2010).

3.5 Population genetic structure

Genetic tools are very useful for tracing the origins of populations (Muirhead et al. 2008; Estoup and Guillemaud 2010). Identifying sources of introduced populations is important for both biodiversity conservation and pest management (Cox 2004; Kenis et al. 2009), as reconstructing colonisation pathways is essential for controlling and preventing future invasions (Wilson et al. 2009).

Genetic diversity plays an important role for the long-term persistence of species. Information on population history, demographic status and the level of gene flow in the landscape can be obtained by using molecular markers, as well as measuring allelic richness within and between populations. Populations founded by a small number of individuals (founder effect) or those that experienced a sudden, massive decline in population size (bottleneck) generally show lower levels of genetic diversity compared to those that frequently exchange migrants and are expanding in population size.

Measuring species dispersal over large temporal and spatial scales using observational methods (e.g. capture-mark-recapture) can be logistically challenging, and even with advanced tracking methods (e.g. radio- and GPS-tracking) it might be difficult to distinguish between dispersal and other types of movement behaviour. Using molecular markers allows measuring the effective dispersal of species, i.e. the realized exchange of genes among populations, which indirectly reflects the functional connectivity of the landscape.

I sampled 29 populations across the entire range of *M. roeselii* in south-central Sweden and genotyped 837 individuals using eight microsatellite markers to identify the source or range expansion and to analyse the patterns of genetic diversity and differentiation in the region (Paper III & IV).
4 Results and Discussion

4.1 Species distribution modelling

Landscape composition is known to affect colonization of habitat patches and the establishment of populations (Hirzel and Le Lay 2008). In the present study (Paper I) the regional occupancy pattern of *M. roeselii* was explained by a number of landscape variables which are linked to habitat suitability and connectivity. Land use types that contain larger amounts of grassland vegetation such as arable land, pastures and rural settlements were identified as the most important variables. But also structural aspects of the landscape such as number of fields and length of linear landscape elements (i.e. roads and streams) were identified to have a considerable influence on the species distribution at the landscape scale.

Testing two different spatial resolutions in the analyses showed that species distribution models at the landscape scale (1 x 1 km) had a greater explanatory power than local scale models (10 m wide buffer strip at both road sides)(Table 1). Models at the landscape scale most likely performed better than local scale models since they not only quantify the amount of available habitat but also take the positive effect of grassy field margins, road sides and ditches on species occurrence into account. This is supported by findings from previous studies which showed that linear landscape elements with grassy vegetation serve as habitat and dispersal corridors for species in agriculture landscapes (Berggren et al. 2001; Meek et al. 2002; Werling and Gratton 2008). The lower performance of models at the local scale can probably be attributed to the coarse resolution of the land use data (25 x 25 m), and the lack of detail of the broad land use categories in the national land cover map (1 : 50 000) which fails to capture subtle differences in local habitat conditions. However, detailed habitat - and species distribution data is seldom available for large spatial extents because it is time consuming to collect (Jimenez-Valverde et al. 2008). In situations when swift action is needed to control invasions or to predict
species distributions across large spatial scales, modelling species distribution using readily available land cover data allows a rapid assessment of the situation (Early et al. 2008; De Groot et al. 2009; Kadoya and Washitani 2010).

Cross-validation showed that the models were generally accurate in predicting species occurrence across the spatial scales for the environmental gradients examined in the study. The models could be used to create a distribution map for *M. roeselii* in south-central Sweden that can be used to identify areas of increased occurrence probability. Such maps can help to plan future surveys, identifying dispersal routes and predicting directions of range expansion in the region.

The present models can be further extended and refined by including variables that account for local population sizes and neighbourhood aspects (number of occupied grid cells surrounding the focal cells) as these factors can influence local occupancy patterns (Kadoya 2009).

**Table 1. AIC weighted model - averaged parameter estimates generated from the top three models (summarized $\omega_i = 0.95$).**

<table>
<thead>
<tr>
<th>Averaged Model</th>
<th>Variable</th>
<th>Coeff</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Landscape Scale</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1 x 1 km)</td>
<td>(Intercept)</td>
<td>-1.478</td>
<td>0.282</td>
</tr>
<tr>
<td></td>
<td>Survey length</td>
<td>1.791</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>Arable land</td>
<td>0.014</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Rural settlements</td>
<td>0.259</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>Pasture</td>
<td>-0.042</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Linear Elements‡</td>
<td>-0.124</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>Fragments†</td>
<td>0.159</td>
<td>0.078</td>
</tr>
<tr>
<td><strong>Local Scale</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10 m buffer)</td>
<td>(Intercept)</td>
<td>-0.996</td>
<td>0.169</td>
</tr>
<tr>
<td></td>
<td>Survey length</td>
<td>0.643</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td>Arable land</td>
<td>1.071</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>Rural settlements</td>
<td>0.559</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td>Pasture</td>
<td>0.100</td>
<td>0.111</td>
</tr>
</tbody>
</table>

### 4.2 Rate of range expansion

Based on the analysis of observational data collected between 1981 and 2010, *M. roeselii* has expanded its range in south-central Sweden (Figure 4) at an estimated speed of 0.31 - 3.16 km/year, depending on the calculation methods
and conceptual models used (area- vs. distance based approach; Table 2). In our study we tested models focusing on different geographic aspects of the range such as the position of the range margin and increase in range size. However, the discrepancy in the estimated expansion rates and the sensitivity of the models to sampling effort (i.e. minimum number of records needed to obtain reliable estimates) was larger than anticipated.

**Figure 4.** Geographical distribution of presence data for *M. roeselii* (mid-point 59°44’N, 16°52’E) in south-central Sweden in 5 x 5 km grid squares (n = 366). Data are gathered from two comprehensive surveys (1989 - 1990, 2008 - 2010) and from the national record database (Species Gateway (1981 - 2010)).

The grid count model (GC) which is based on measuring the change in occupied area over time returned the lowest rate of range expansion (0.31 km/year), while rates of range expansion estimated using distance-based methods were considerably higher (Table 2). Distance-based models are most suitable for detecting positional shifts of the range margin, as they measure the distances between the source of range expansion and the observations from consecutive years. According to the Median (MdN)- and Mean (M)-model *M. roeselii* has been expanding in its main distribution at rate of 0.93 km/year and 0.97 km/year respectively. Estimated rates of range expansion were even higher for models using exclusively observational data from the outer range margin (1.45 - 3.16 km/year; conditional maximum (CMX)-, margin (MR)-, maximum (MX)-, and 95th gamma quantile (GAM)-model; Table 2).
Table 2. Estimated rates of range expansion (km/yr) derived from different range statistics using all available data (1981-2012; n = 366) and (2) pruned datasets based on the minimum number of occupied sites needed to be surveyed to achieve 90% accuracy of the range statistic (1989-1990 & 2008-2010; n = 341). The percentage and actual numbers of occupied squares needing to be sampled on a yearly basis to give 90% accuracy to the range statistics are given in column 3. Percent changes in estimates are the differences between the expansion rates calculated using the pruned dataset relative to the whole dataset.

<table>
<thead>
<tr>
<th>Range expansion model type</th>
<th>Expansion rate (all data)</th>
<th>Min. sample for 90% accuracy</th>
<th>Expansion rate (pruned data)</th>
<th>Percent change in estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid count (GC)</td>
<td>0.31</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean (M)</td>
<td>0.97</td>
<td>36</td>
<td>1.52</td>
<td>+56</td>
</tr>
<tr>
<td>Median (MdN)</td>
<td>0.93</td>
<td>90</td>
<td>1.50</td>
<td>+61</td>
</tr>
<tr>
<td>Gamma quantile (GAM)</td>
<td>3.16*</td>
<td>22</td>
<td>3.05</td>
<td>-3</td>
</tr>
<tr>
<td>Marginal (MR)</td>
<td>1.72</td>
<td>41</td>
<td>2.56</td>
<td>+49</td>
</tr>
<tr>
<td>Maximum (MX)</td>
<td>2.39</td>
<td>17</td>
<td>3.09</td>
<td>+33</td>
</tr>
<tr>
<td>Conditional maximum (CMX)</td>
<td>1.45</td>
<td>17</td>
<td>3.09</td>
<td>+113</td>
</tr>
</tbody>
</table>

1. The minimum number of occupied sites needing to be sampled was based on the 1989-1990 survey
2. The 95th Gamma quantile could not be calculated for years 1981, 1986, 2003 and 2005 due to a single record in those years.

The area-based GC-model predicted the smallest rate of range expansion as it assumes that the species disperses according to a simple diffusion model (Van den Bosch et al. 1990; Lensink 1997). To obtain reliable estimates on the rate of range expansion using the GC-model, an extensive sampling across the entire distribution range at regular time intervals is required. In our case sampling effort (intensity, extent) was highly variable since the available occupancy data originated from multiple sources (Species Gateway; surveys 1989-1990 and 2008-2010). Thus, inclusion of data from years in which the area occupied by the species was largely under-sampled would lead to an underestimation of the actual rate of range expansion in *M. roeselii*.

The MdN-model as well as the MR- and M- model likewise are sensitive to sampling effort and require a larger number of annual observations to calculate range expansion rates with 90% accuracy (Table 2). Both the M- and MdN-model might therefore fit less well for estimating the range expansion of species like *M. roeselii* that display two different kinds of dispersal behaviours, slow continuous short-distance dispersal and infrequent long-distance dispersal. As the few long distance values have relatively little influence on the mean and median estimates, the M- and MdN-models may thus underestimate the speed of range expansion.
Marginal models (MR-, CMX-, MX-, GAM-model) on the other hand measure the expansion rate based on rare long-distance dispersal events and only few observations are needed to obtain statistically confident estimates (Table 2). These models give a good approximation for the establishment rate of pioneer populations at the forefront of the main range margin. Consequently such estimates are of great importance for developing conservation and control strategies, as pioneer populations lead to sudden and rapid increases in the occupied area while at the same time being easier to eradicate than large established populations in the core of the range (Moody and Mack 1988; Hulme 2003; Kovacs et al. 2011).

The observed discrepancy in our estimated expansion rates could result from separate processes occurring in different parts of the species range (Simmons and Thomas 2004). Range boundaries are sensitive to stochastic fluctuations that frequently occur in marginal populations (Brown et al. 1996) (Berggren et al. 2009). Marginal areas are likely to be colonised by individuals with high dispersal capability, but dispersal propensity tends to drop within only few generations after establishment due to the trade-off between reproduction rate and dispersal ability (Zera and Denno 1997; Hochkirch and Damerau 2009). This effect is supported by observations in other Orthopterans, for example are the predicted rates of spread for long-winged *Conocephalus discolor* nearly six times larger at range margins compared to those at the core of the range (6.70 vs. 1.24 km/year; Simmons 2003).

Sensitivity analysis showed that the number of annual records of *M. roeselii* in the national database (Species Gateway) was too low to be used by itself for the estimation of the rate of range expansion. However, in combination with comprehensive surveys these data are valuable as they increase sample size and recording effort and indicate new locations for structured surveys. For certain species that are rapidly expanding their range while at the same time being difficult to detect, it is a challenge to monitor their distributions over large spatial extents. Since organized large-scale surveys at regular time intervals are expensive in cost and time, citizen collected data can be highly valuable to estimate species distributions (Goffredo et al. 2010; Snäll et al. 2011).

### 4.3 Local origin and genetic patterns across the range

Based on historic trade records and first sightings of the species it was previously assumed that *M. roeselii* had arrived with international shipping cargo in the inland harbour city of *Västerås* and that its range expansion started from there (Figure 5). Our analyses of the genetic diversity of 29 populations sampled across the species range in the Lake Mälaren region revealed that
allelic richness was highest in the Ullvi population situated 15 km (aerial distance) south-west of Västerås. Thus the area of range expansion most likely started somewhere around Ullvi. The Royal estate of Strömsholm just 5 km east of Ullvi was home to the Royal Horse Academy since the 17th century and historic documents confirm repeated import of cattle and horses from continental Europe for breeding programs (Bäckström 1924). Metrioptera roeselii may thus have arrived in the forage material (hay) with the imported animals, which is further supported by previous genetic analyses showing German or Polish ancestry of M. roeselii in the Mälaren region (Kaňuch et al. unpublished).

Figure 5. Geographic location and genetic structure of the 29 sampled populations in the Mälaren region (59°35’N, 16°30’E), south central Sweden. The site number corresponds to the Population identifier (Pop ID). ID number 29 = Västerås, ID number 8 = Ullvi. The colour of the squares depicts the ten genetic clusters identified by BAPS. The thickness of the squares outline illustrates the two genetic clusters identified by STRUCTURE.
Genetic diversity was on average higher in populations sampled in the core area of the range compared to those at the range margin (Figure 6). This matches theoretical predictions as genetic diversity is known to be positively related to population age, population size and habitat continuity (Haag et al. 2005; Holzhauer et al. 2009). Previous records confirm the continuous presence of M. roeselii in the area around Västerås and Lake Mälaren since 1970 and its increase in abundance in the last four decades (Idar 1973; de Jong and Kindvall 1991; Species Gateway 2010). Moreover, we have shown (Paper I) that the occurrence of M. roeselii can be predicted by the proportion of arable land, which constitutes the dominant land use type in the Lake Mälaren region. Therefore our findings of high genetic diversity and levels of gene flow in the centre of the species range match the predictions that local population size and frequency of dispersal are positively related to habitat continuity.

Figure 6. Genetic diversity in populations of M. roeselii sampled at an increasing distance from centre of distribution (Ullvi): a) Allelic richness (AR), b) Private Allelic richness (PAR), c) Unbiased gene diversity (UHe), and d) Inbreeding coefficient (Fis).
Frequent dispersal also explains the low genetic differentiation between populations in the core of the species range. Populations at the range margin, on the other hand, showed reduced allelic richness and a higher degree of differentiation, not only between themselves but also in comparison to populations in the core of the range. This suggests that the marginal populations were recently founded by a smaller number of individuals that dispersed over long distances either actively by flight and/or passively through human-mediated dispersal. The fact that we did not find any evidence of recent bottlenecks in any of the sampled populations and only slightly increased levels of homozygosity indicate that gene flow was sufficiently high to reduce founder effects via random mating and migration. The gradually decreasing levels of genetic diversity in relation to increasing distance from the core area and evidence of a weak but significant isolation-by-distance pattern (Figure 7) could be explained by the species dispersal behaviour of combined frequent short distance dispersal and occasional long-distance events. Based on 837 individuals from 29 populations, BAPS identified ten genetic clusters in the study region (Figure 5). Populations in the core of the range form two genetic clusters whereas marginal populations split into eight clusters, suggesting that gene flow is limited between marginal populations as well as between core and marginal populations. However, the recently established populations at the range margin have presumably not yet reached a demographic equilibrium and the genetic composition is therefore expected to change with time as a consequence of genetic drift, mutations and gene flow (Bohonak 1999).

The amount of gene flow reflects not only the species dispersal behaviour but is also influenced by the characteristics of the landscape. With increasing distances from the Lake Mälaren, the landscape structure and composition becomes more mosaic-like consisting of arable land, forests and small lakes. Previous studies (Hunter 2002; Dyck and Baguette 2005; Merckx and Van Dyck 2007) have shown the negative effect of habitat fragmentation on species dispersal. The extent to which habitat fragmentation hinders species dispersal depends on the severity of fragmentation, the spatial arrangement of habitat patches and the species dispersal ability (Diekötter et al. 2007; Heidinger et al. 2010). In Paper IV we explore the influence of the landscape on dispersal by measuring the functional connectivity in a subsample of ten populations and relating to measures of genetic diversity and differentiation.
4.4 Functional connectivity-dispersal routes

Least-cost paths (LCP) through forested landscapes were twice to three times longer than the Euclidian distance between populations, whereas in arable landscapes near Lake Mälaren least-cost paths were 1.1-1.3 times longer than the Euclidian distance (Figure 8 a and b). The least-cost paths between populations explained less of the genetic differentiation (Fst value) than the Euclidian distance. This either means that landscape features influencing species dispersal were not included correctly in the LCP analysis or that species dispersal is not entirely confined to movement through suitable habitat in the classification used in the analysis. Our results indicate that long winged individuals and unintentional human translocation of individuals could enable the species to bypass less suitable habitats. This explains why least-cost path analyses proved to be inconclusive in this case. It thus would be interesting if improving the LCP analyses by looking at more closely located populations or populations in landscapes with different degrees of habitat fragmentation. Moreover, by assigning more differentiated resistance weights to land use covers to change the contrast between habitat types would give a more realistic picture of species dispersal routes.

The relationships between allelic richness and unbiased gene diversity were not significantly correlated with either road density or number of farms around the populations (Figure 9 and 10). We did observe a tendency towards
increased genetic variation suggesting that both farms and road density improve the connectivity of habitat patches and thereby have a suggestive positive influence on the species’ dispersal. Higher connectivity should lead to populations being exposed to higher migration rates resulting in an increase in effective population size (Cushman et al. 2006). Moreover, larger effective population sizes are expected to increase allelic richness and unbiased gene diversity of populations (Aspi et al. 2006).

Figure 8a. The ten study populations of *M. roeselii* in south-central Sweden (midpoint 59°85’N, 16°25’E). GSD land cover map before reclassification for least-cost analyses.
**Figure 8b.** Calculated least-cost paths between study populations of *M. roeselii* in south-central Sweden. Habitat classifications were based on the Swedish GSD Land cover map. Habitats are given different shades of green in declining order of species preference: arable land, pastures, rural settlements (weight 1 - 3), orange: mainly forest and urban areas (weight 4) and blue: water bodies (weight 5).

**Figure 9.** The density of roads in 5 x 5 km squares around the study populations of *M. roeselii* and the genetic characteristics of the different populations: a) AR = 6.4635 + 0.0071 x road density and b) UHe = 0.6143 + 0.0007 x road density.
Figure 10. The density of farms in 5 x 5 km squares around the study populations of *M. roeselii* and the genetic characteristics of the different populations: a) $AR = 6.5896 + 0.0142 \times$ farm density and b) $UHe = 0.6232 + 0.0016 \times$ farm density.
5 Conclusions

My thesis shows that the regional distribution of *Metrioptera roeselii* can be predicted by a number of landscape variables: the amount of arable land, pasture, rural settlements and linear landscape elements (Paper I). The possibility to model the species distribution using survey data and available land-cover data is promising in that it will enable us to predict the direction and possible extent of future range expansion of the species. Results from this study may also be valid for species with similar habitat preferences and dispersal behaviour. Predictions of occupancy patterns could be further improved by incorporating demographic parameters (e.g. local abundance) and neighbourhood aspects (e.g. distance to nearest population) into the distribution model.

Another important aspect of range expansions is the velocity of spread. For the first time the rate of range expansion of *M. roeselii* in south-central Sweden has been estimated. The obtained estimates ranged from 0.3 - 3.2 km/year depending on the type of statistical method used (Paper II). These estimates assume a constant rate of spread in a homogenous environment. However, the landscape in the study region is heterogeneous and the rate of range expansion is therefore likely to vary depending on the continuity and connectivity of habitats. In this regard range expansion models could be improved by including a measure of landscape heterogeneity to describe more realistically observed patterns of range expansion.

By combining population genetics with landscape analyses, we found that *M. roeselii* is likely to disperse efficiently via three dispersal modes: common short distance dispersal, occasional long distance dispersal or unintentional human-mediated translocation. Genetic diversity in recently founded populations at the range margin was only slightly reduced which suggests a frequent influx of migrants. The weak correlation between geographic distance measures and genetic distance shows that dispersal is neither limited to
neighbouring populations nor confined to movement through suitable habitat (Paper III). This may explain why least-cost path analyses proved to be inconclusive in our case (Paper IV). Long-distance and human aided dispersal appear to occur more frequently than previously assumed. The ability to use multiple dispersal pathways and to colonize a range of grassland types enables the species to spread effectively through fragmented landscapes. However, further investigations at a finer spatial resolution, testing different land use classifications and cost-weightings of habitat types would be useful to further increase our knowledge on the effect of the landscape on patterns of gene flow.
References


Species Gateway 2010. Swedish Species Information Centre, Swedish Environmental Protection Agency.


Acknowledgements

I like to thank Nicole A Schneider, Abhi Ratnakumar, Tytti Vanhala, Juan C Ruiz Guajardo, and my supervisors for valuable comments on my thesis. I also like to thank Celina Abraham, Lina Ahlbäck, Emilia Boberg, and Carin Erikson for helping in the field, Johan Wallén for technical advice in the genetics lab. The work was funded by Formas, Carl Trygger’s foundation and Oscar and Lili Lamm’s foundation.
Tack!

This work would have not been possible without the help and support from my colleagues, friends and family.

I especially like to thank:

Åsa Berggren and Anna Cassel-Lundhagen for facilitating the development of interesting projects and giving me the opportunity to learn a broad range of valuable scientific tools.

Matt Low for statistical advice and proof reading of two of my manuscripts.

Göran Nordlander for general comments on the last chapters of the thesis.

Emilia B., Lina A. and Carin E. for assistance in the field.

Peter K. for getting me fascinated about grasshoppers and crickets.

Celina for being the most enthusiastic field assistant, making catching hundreds of crickets fun.

Alejandro for help with statistical & graphical matters and cups of mate.

Per N. for solving computer issues in no-time and for inspiring conversations.

Tytti for teaching me about population genetics, providing help with the data analyses and valuable comments on the genetic chapters of my thesis.

Katja for her patience and support when things didn’t work out the way I had hoped for.
Abhi for sharing PhD worries and laughing about them over a glass of wine.

JC for encouragement and advice during the last years.

Karin P. for being the sunshine on a cloudy day.

Vera and Lel for uplifting conversations.

Peter for getting me back on track.

Robin for making me laugh.

Nicole for the fantastic help and support right to the end.

Sara for being an extra mum and valued friend.

Christel for getting me interested in biology from an early age.

Olaf for being the best brother ever.

Mum & Dad for being fantastic parents, encouraging and supporting no matter what.