

Original article

# Estimation of habitat quality based on plant community, and effects of isolation in a network of butterfly habitat patches

Javier Sawchik<sup>a,b,\*</sup>, Marc Dufrière<sup>a,c</sup>, Philippe Lebrun<sup>a</sup>

<sup>a</sup> *Unité d'écologie et de biogéographie, Centre de recherche de la biodiversité, Université catholique de Louvain, Louvain-la-Neuve, Belgium*

<sup>b</sup> *Sección Ecología Terrestre, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay*

<sup>c</sup> *Ministère de la région Wallonne, Centre de recherche de la nature, des forêts et du bois, Observatoire de la Faune, de la Flore et des Habitats, Belgium*

Received 26 November 2001; received in revised form 28 June 2002; accepted 7 November 2002

## Abstract

Minimum viable site networks are crucial for many threatened species but their design is always a difficult procedure. The present study investigated methods to estimate habitat quality of patches, in an ecological network for five butterfly species (*Brenthis ino*, *Clossiana selene*, *Lycaena helle*, *Lycaena hippothoe*, *Procllossiana eunomia*) inhabiting wet meadows. Abundance predictions were performed on the basis of botanical relevés with a multiple-species approach combining canonical correspondence analysis and multiple regression. Models are defined on a reference site set and are evaluated on a test site set. All the fitted models predicted the abundance of the butterfly species considerably well (from 61 to 86% of the variation). We evaluated the potential consequences of isolation on local populations, by comparing predicted and observed abundance. It was expected that greater differences would be observed when sites were more isolated. On the test set, differences between predicted and observed abundance were largely correlated to site isolation for *L. helle* and *P. eunomia*. The most isolated sites had greater chances to be empty, even if they had high-quality habitat. Reciprocally, when the sites were less isolated, the abundance observed was always greater than predicted, seeming to confirm the role of rescue effects.

© 2003 Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

**Keywords:** Butterfly assemblage; CCA; Habitat connectivity; Habitat quality; Isolation effect; Predictive model

## 1. Introduction

With 5% to 10% of extinct species and 30% to 50% of decreasing species in the last century, southern Belgium and its surrounding regions are among the highest impacted areas by human activities in Middle Europe (EEW, 1993). The main reason for these figures is habitat destruction and fragmentation that result from a high human population density (300 people per km<sup>2</sup>) and very intensive agriculture and forestry. In such perturbed and productivity-oriented environments, it is important to assess key sites that must be protected and sites that should be extensively managed or restored.

Many butterfly species living in fragmented landscapes show spatially structured populations that may fit to the metapopulation concept (Baguette and Nève, 1994; Hanski, 1998; Fischer et al., 1999). Metapopulation dynamics are

regulated by the processes of local extinction and (re)colonisation (Hanski, 1998). To ensure the effective conservation of threatened species, it is fundamental to preserve habitat patch networks. Generally, it is not possible to preserve all patches in a network, so we have to decide which patches should be protected with priority. The fact that a patch is empty does not necessarily indicate that it is not suitable. It could be temporarily empty as a consequence of extinction-colonisation stochasticity (Hanski, 1998). Generally, the most isolated patches should have low probabilities of being occupied. In order to take objective decisions, it is necessary to develop methods that allow the estimation of the potential of habitat patches in a network. This measure will depend on the area, the habitat quality and on the connectivity of habitat patches (Thomas et al., 2001), and may be used in the development of spatially explicit models.

Spatially extended population viability analysis seems to be the best solution to identify key sites in a network (Shaffer, 1990; Boyce, 1992; Lindenmayer et al., 1995; Burgman et al., 1996; Hanski, 1998; Moilanen, 1998). Spatially realistic

\* Corresponding author.

E-mail address: [sawchik@fcien.edu.uy](mailto:sawchik@fcien.edu.uy) (J. Sawchik)

models require the development of measures of habitat quality in order to estimate the carrying capacity or the effective area of potential habitat patches (Moilanen and Hanski, 1998). Unfortunately, some of these models require accurate estimations of several biological parameters, which are not easily measurable in invertebrate organisms because of their small size and complex life cycles. Other(s) approaches are based on indices of habitat suitability that are derived from diverse characteristics of the sites, e.g. flower abundance and richness, microtopography, vegetation structure or Ellenberg indicator values. These indices may be used to predict the presence or abundance of butterflies through canonical ordination, discriminant function analysis or generalised linear models (Kremen, 1992; Oostermeijer and van Swaay, 1998; Thomas et al., 2001).

In a first approach, we can assume that local density of any species at a given site is constrained by site area and quality. So, a small set of ecological parameters could be used to predict the potential of sites to support local populations of threatened species. Hence, it should be possible to identify sites with very high potential but not actually colonised or occupied by a few specimens. The identification of such weakness in a network is very useful to precise local problems to better understand the population system or to identify other limiting factors.

This paper aims to evaluate the usefulness of botanical relevés as a measure of habitat quality to estimate the abundance of threatened butterfly species. A botanical relevé was defined as a vegetation sample that describes the floristic composition of a quadrat (Kent and Coker, 1992). Experienced naturalists are usually able to predict the presence and frequency of many species in a site only from a quick glimpse. It is those field perceptions that we intend to understand and model through a botanical relevé. We propose to use plant community as the variable to describe the quality of the habitat patches. One advantage of this choice is that plants are the essential resource for butterfly survival, providing nutrients for caterpillars and adults. Caterpillars are intimately associated with their host plants (Gilbert and Singer, 1975; Chew, 1980; Porter, 1992; Ravenscroft, 1994). Likewise, adult individuals need some plant species as nectar sources (Murphy, 1983; Shreeve, 1992; Holl, 1995; Lortscher et al., 1995). Furthermore, the vegetation responds to a large range of environmental and historical factors, e.g. altitude, temperature, precipitation, topography, land use, soil type, microclimates (Ellenberg et al., 1992). Indeed, vegetation is also the main physical interface between butterflies and their environment. Thus, a strong relationship between butterfly species and plant communities is expected.

Five species of diurnal butterfly species (*Brenthis ino*, *Clossiana selene*, *Lycaena helle*, *Lycaena hippothoe* and *Proclissiana eunomia*) were chosen as flagship species to set up an operational ecological network. They are all threatened species in southern Belgium (Goffart et al., 1992).

Finally, we discuss the extent to which patch occupancy can be predicted from the combination of patch isolation and

habitat quality. It is expected that greater differences between predicted and observed abundances will be observed when the sites are more isolated.

The main objectives of this work were: (1) to develop a method to predict butterfly abundances based on vegetation data, (2) to assess the importance of nonhost plants as predictors of butterflies in comparison to host plants, and (3) to evaluate the isolation effect on the expected abundances predicted by the model.

## 2. Materials and methods

### 2.1. Study area

The Plateau des Tailles is a high plateau (max. 654 m) located in the Ardenne biogeographical area, southern Belgium. In this area, there are many semi-natural habitats like peat bog, fen and moor relicts already protected as nature reserves. Several valleys furrow this high plateau, descending below 200 m, where there are many highly interesting, semi-natural meadows still not protected. These meadows are not well phytosociologically defined because in many cases they include several kinds of very complex and highly patchy humid grassland plant communities. However, several of these habitat categories are mentioned in the first annex of the 92/43/CEE Fauna-Flora-Habitat Directive as habitats that should be protected.

The study area corresponds to the oro-hydrographic basin of the Martin-Moulin stream, located in the Plateau de Tailles (Fig. 1). A large range of meadow types are represented in this area, from wet hay meadows at lower altitude to *Juncus* spp. and *Carex* spp. dominated fens in the higher locations.

### 2.2. Sampling

The potential site network included 34 sites (reference set) where vegetation and butterfly populations were surveyed. In

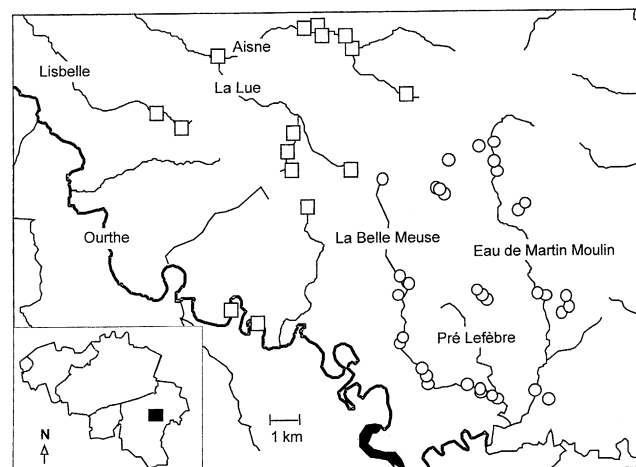


Fig. 1. Studied area in the Plateau des Tailles, Ardenne, Belgium. Reference sites are represented by circles and test sites by squares.

addition, 13 sites (test set) located in the neighbouring Aisne and Ourthe valleys were selected to validate the models and to assess their predictive power. Field surveys in surrounding areas demonstrated that the test sites were more isolated than the reference sites. Test sites are located on the west margin of the distribution area of *C. selene*, *L. helle*, *L. hippothoe* and *P. eunomia*. Moreover, the closest populations to the north and the south were located more than 8 to 10 km away.

In every site, 10 edaphic variables were measured in May 1996. Five soil core samples were randomly taken in each site, and subsequently pooled. These samples were analysed for moisture content, pH (H<sub>2</sub>O), pH (KCl), organic carbon, potassium, phosphorus, sodium, magnesium and calcium. The depth of the water table was recorded in three locations at each site and average values were computed to be used in subsequent analyses.

Vegetation surveys were conducted from May to July in 1995 and in 1996. After being cautiously inspected by an experienced botanist, the sites were defined as homogeneous botanical associations. Site boundaries were delineated on the basis of changes in floristic composition. Heterogeneous sites were subdivided into smaller patches showing low internal floristic variability (Daiyuan et al., 1998). At each site, vegetation was sampled by a detailed species relevé. Cover percentage of vascular plant species was estimated in a random located quadrat of 5 × 5 m (Kent and Coker, 1992). Then, the size of the quadrat was progressively increased until no new species were found. Cover values were averaged over the two sampling periods and multiplied by the site area. A total of 149 plant species were recorded over all the sampled sites. Rare species, defined as those plant species that were not present in at least four sites (approximately 10% of the total), were ignored for subsequent analyses. A vegetation matrix was constructed comprising the distribution of 63 plant species in the 34 sampled sites.

We sampled butterflies in each site as many times as possible during the adult flight period, i.e. May–August, between 1994 and 1996. Each site was visited at least three times per year during the flying period of each butterfly species. We proceeded by visual observation and/or collection of adult individuals along zigzag transects during fine weather conditions (Pollard, 1977; Thomas, 1984). All the observations were done by the same team (three persons) to avoid collector effects.

In order to control sampling effort differences, we used the maximum abundance of each species in each site as a measure of relative abundance (Smallidge et al., 1996). Field experience in such habitats demonstrated that after three or four well distributed visits during the main activity period that measure does not change significantly. A preliminary comparison with Mark Release Recapture (MRR) results (Baguette and Nève, 1994; Nève et al., 1996) indicates that maximum abundance is a good estimation of the relative abundance. Data from the years 1994–1996 were pooled to avoid annual weather effects that could modify the number of specimens observed, as well as local extinction effects. The

butterfly matrix shows the abundance of the five butterfly species in the 34 sampled sites.

### 2.3. Statistical analyses

#### 2.3.1. Butterfly-environment relationships

Spearman rank coefficients were computed to assess the correlation between plant and butterfly species. Nonparametric rank coefficients are robust to outliers and less sensitive than Pearson's product-moment correlation coefficients to high number of zeros in the data matrix (Potvin and Roff, 1993; Legendre and Legendre, 1998).

Detrended correspondence analysis (DCA) (detrending by segments) was performed on the butterfly matrix to evaluate the length of the main gradient. The gradient length, which is measured in standard deviation units of species turnover (SD), allows to select between linear or nonlinear methods (Jongman et al., 1995). The DCA showed that the first axis had a gradient length over 3 SD, suggesting that species responses may be moderately unimodal (Jongman et al., 1995).

Butterfly-environment relations were inferred using CCA, which maximises the dispersion of species scores assuming approximately unimodal response models (ter Braak, 1986; ter Braak and Prentice, 1989; Palmer, 1993). Since the ordination axes are constrained to be a linear combination of ecological variables, CCA is a technique for direct gradient analysis (ter Braak, 1986). However, if the number of explaining variables is the same as the number of the studied sites, CCA becomes an unconstrained analysis (ter Braak, 1986). Thus, the first step in order to explore the relationships between butterfly species and plant species with CCA is to reduce the number of potential explaining variables (63 plant species) into a small and informative set. First, we introduced in the model the host plant species (*Filipendula ulmaria*, *Polygonum bistorta*, *Rumex acetosa*, *Viola palustris*). Second, we performed stepwise CCA with the 59 other plant species as explaining variables. The variables were selected manually on the basis of additional variation explained, and their significance was tested by a Monte Carlo permutation test (199 permutations). Finally, we performed CCA on the butterfly matrix and a set of 10 plant species as explaining variables. The number of explaining variables was fixed as 10 in order to have comparable results with the CCA conducted using 10 edaphic factors as environmental variables. An additional CCA was run using the four host plants only as explaining variables. Monte Carlo permutation tests (199 permutations) were used to assess the significance of the first two canonical axes showing the relationship between butterfly and plant species (ter Braak, 1988; ter Braak and Smlauer, 1998).

CCA was also performed on the butterfly data using the edaphic factors as explaining variables. The fraction of variation in the butterfly data explained by the vegetation data independent of the edaphic factors was determined through a series of partial CCAs (Jongman et al., 1995). Partial CCA is

a modification of CCA that produces an ordination of the species data constrained by one set of environmental variables while controlling for the effect of a second set of environmental variables named covariables.

### 2.3.2. Spatial structure of the data

Spatial autocorrelation is a common phenomenon in ecological studies (Legendre and Fortin, 1989; Legendre, 1993). We removed spatial effects from the butterfly data set using partial CCA and the variation partitioning method (Borcard et al., 1992). The terms of a third-degree surface trend polynomial equation of the geographical coordinates ( $x$ ,  $y$ ) of the sites were introduced in the model and significant terms were selected by the forward selection option (ter Braak and Smlauer, 1998; ter Braak, 1990). We analysed spatial structure individually for each butterfly species using autocorrelograms (with Moran's  $I$ ) computed with 10 equidistant distance classes (Legendre and Legendre, 1998). The distance unit was circa 1 km.

### 2.3.3. Predicting butterfly abundance

Canonical coefficients are the coefficients of a weighted multiple regression of the site scores on the standardised environmental variables. Such coefficients allow the location of the sites in the species-space according to the environmental variables, but they are not sufficient to predict the species abundance in those sites.

Stepwise multiple regression was used to model the abundance of the butterfly species as functions of the site scores in canonical axes. In this way, we used stepwise CCA as exploratory, as well as a variable reduction tool. We also considered quadratic expressions of canonical scores, in order to account for possible unimodal relationships between species and ecological factors (Jongman et al., 1995). Butterfly abundances were log-transformed when necessary for regression assumptions. For stepwise procedures,  $P = 0.05$  and  $P = 0.10$  were used as entry and exit threshold values, respectively (Post and Stenseth, 1998). Model assumptions were checked by visual inspection of residual plots (Dowdy and Wearden, 1985). This multiple-species approach (CCA combined with regression on canonical axes) seems to give better predictions compared to single-species regression approaches (Sawchik, 1999). One possible explanation is that CCA selects the plant species, giving priority to those that explain the abundance distribution of several butterfly species.

The canonical coefficients obtained with the CCA performed on the reference sites were used to compute the site scores of the test sites. After being standardised, the plant frequencies on the test sites were multiplied by the canonical coefficients and summed to give the site score in the butterfly species space. The site scores on the canonical axes were introduced in the regression models to obtain estimates of the butterfly abundances on the test sites. We computed Pearson's product-moment correlation coefficients between predicted and observed abundance of the butterfly species. Data were not bivariately normal; so, significance tests of the

correlation coefficients were performed with permutation procedures (Legendre and Legendre, 1998).

### 2.3.4. Measuring the significance of site isolation

The test set comprised a lesser number of sites distributed over a similar surface area and then with a lower connectivity between the sites. Under isolated conditions, we could expect greatest differences between predicted and observed abundances. We computed the average distance to the five closer occupied sites as a measure of site isolation, for each species. All the information available regarding species presence on the Plateau des Tailles was used to compute isolation indices. We determined the correlation between isolation and model residuals, i.e. differences between observed and predicted abundances, by using Pearson's product-moment coefficients.

DCA, CCA, stepwise CCA and partial CCA were carried out using CANOCO (ter Braak, 1988; ter Braak and Smlauer, 1998). The R package (Legendre and Vaudor, 1991) was used to compute correlation coefficients and autocorrelograms. SAS software (SAS, 1982) was used to perform stepwise regressions.

## 3. Results

Relative abundances of the five butterfly species in the reference site set are summarised in Table 1. The five butterfly species occurred in a relatively high number of sites showing occupancy levels between 50% (*C. selene*) and 76% (*B. ino*).

Table 2 synthesises the best five Spearman rank correlations between plant and butterfly species. Each butterfly species was significantly correlated with its host plant and with some other plant species, as well. When CCA used exclusively host plants, it explained 49% of the variation in the butterfly data and it succeeded in explaining 76% of the variation when 10 plant species were used.

The 10 plant species selected by stepwise CCA were *Carex ovalis*, *Crepis paludosa*, *Festuca rubra*, *Filipendula ulmaria*, *Galeopsis tetrahit*, *Molinia coerulea*, *Polygonum bistorta*, *Potentilla erecta*, *Rumex acetosa* and *Viola palustris*. All variables in the model were significant at the 0.10 level.

The first four canonical axes explained 76% of the butterfly data set variation. We concluded that the 10 plant species explained the main patterns of variation in butterfly species composition. The variance inflation factors were below 10 for all variables considered, indicating that multicollinearity

Table 1  
Average relative abundance per site, range, and number of occupied sites of each butterfly species in the reference site set

	<i>B. ino</i>	<i>C. selene</i>	<i>L. helle</i>	<i>L. hippothoe</i>	<i>P. eunomia?</i>
Average	23.5	5.7	.1	3.0	6.5
Range	0–92	0–53	0–38	0–15	0–35
Occupancy	26	17	22	18	22

Table 2

Spearman correlation coefficients between butterfly abundance and plant frequencies. Host plant species are indicated in bold face

<i>B. ino</i>	<i>C. selene</i>	<i>L. helle</i>	<i>L. hippothoe</i>	<i>P. eunomia</i>
<b><i>Filipendula ulmaria</i></b>	<i>Filipendula ulmaria</i>	<b><i>Polygonum bistorta</i></b>	<i>Festuca rubra</i>	<b><i>Polygonum bistorta</i></b>
0.719	-0.635	0.597	0.609	0.604
<i>Agrostis canina</i>	<i>Urtica dioica</i>	<i>Ranunculus flammula</i>	<i>Crepis paludosa</i>	<i>Ranunculus flammula</i>
-0.579	-0.554	-0.590	0.514	-0.433
<i>Galium saxatile</i>	<i>Agrostis canina</i>	<i>Juncus acutiflorus</i>	<b><i>Rumex acetosa</i></b>	<i>Carex ovalis</i>
-0.558	0.549	-0.558	0.554	0.399
<i>Holcus lanatus</i>	<b><i>Viola palustris</i></b>	<i>Ajuga reptans</i>	<i>Scirpus sylvaticus</i>	<i>Myosotis nemorosa</i>
-0.442	0.546	-0.512	-0.463	-0.351
<i>Valeriana repens</i>	<i>Potentilla erecta</i>	<i>Stellaria graminea</i>	<i>Galium aparine</i>	<i>Deschampsia cespitosa</i>
0.401	0.542	0.434	-0.451	0.336

effects were negligible (ter Braak, 1990; ter Braak and Smilauer, 1998). No outliers, i.e. samples with extreme values in the explanatory variables, were found. Since the axes did not show similar eigenvalues, we concluded that the ordination axes were not unstable (Øksanen and Minchin, 1997; ter Braak and Smilauer, 1998). *Clossiana selene* and *B. ino* were the best explained species (82% and 77%, respectively). *Lycaena helle* and *P. eunomia* followed, reaching 74% and 72% of explained variation, respectively. Abundance variation of *L. hippothoe* was least associated with plant frequencies (63%).

The first two axes explained 70% of the butterfly data set variation and 95% of the species-environment (butterfly-plant) relationship. These axes showed a species-environment correlation of 0.913 and 0.879, respectively, indicating a very high similarity between the site scores computed from botanical relevés and those from butterfly abundances. The first and second canonical axes were significant at the 0.01 level.

The biplot of Fig. 2 captures 70% of the total butterfly variation and correctly identifies the main ecological factors. The first canonical axis isolated *C. selene* and *L. hippothoe* to a lesser extent; the second axis opposed *L. helle* and *P. eunomia* to *B. ino*. *Viola palustris* was clearly associated with *C. selene*, in the first axis (81% of explained variation). *Polygonum bistorta*, *C. ovalis* and *F. rubra* pointed in the direction of *L. helle* and *P. eunomia*. *Brenthis ino* was also well correlated with its host plant *F. ulmaria*. The second axis explained more than a half of the variation of *P. eunomia* and *L. helle* and 25% of the variation of *B. ino*.

On the third axis (Fig. 3), *L. hippothoe* was associated with *R. acetosa*, its host plant, and *C. paludosa*. The last axis associated *P. eunomia* to *F. rubra*, *M. coerulea* and *P. erecta* opposed to *G. tetrahit*.

Edaphic factors explained only 24% of the butterfly variation compared to the 76% explained by vegetation data. Partial CCA showed that 53% of the variation was explained by vegetation data only, when the shared variation with the edaphic factors was removed.

The only significant spatial variable selected by the forward selection procedure was the ( $x \times y$ ) interaction, which reflects a latitudinal-longitudinal trend. The partial CCA indicated that the variation explained by the 10 plant species

(76%) was greater than that explained by the spatial structure (26%). Moreover, the variation explained by the plant species, independent of any spatial structure, was still high (52%). These results are summarised in Table 3. Only *L. hippothoe*, and to a much lesser extent, *C. selene*, showed a strong autocorrelated pattern. Significant autocorrelation values ( $P = 0.001$ ) were positive for the first distance classes and negative for the last distance classes.

Results of stepwise multiple regressions with butterfly abundances as response variables and canonical scores (linear and quadratic expressions of the first four axes) as predictor variables are synthesised in Table 4. In spite of the low scores, the third and fourth axes were retained because they explained 24% and 14% of *L. hippothoe* and *P. eunomia* variation, respectively. Canonical axes were not significantly correlated between them, so no multicollinearity effects were suspected. Model assumptions were highly improved by log-transformation for *B. ino* and *L. hippothoe*. Quadratic expressions of canonical scores were not significant for *L. hip-*

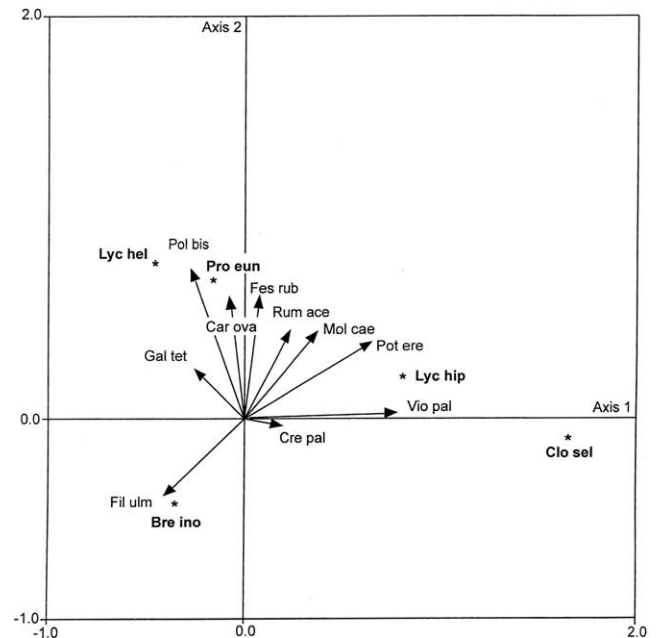


Fig. 2. CCA ordination biplot of the butterfly and plant species relations on axes 1 and 2. Butterfly and plant species names are abbreviated. Butterfly species are indicated in bold face.

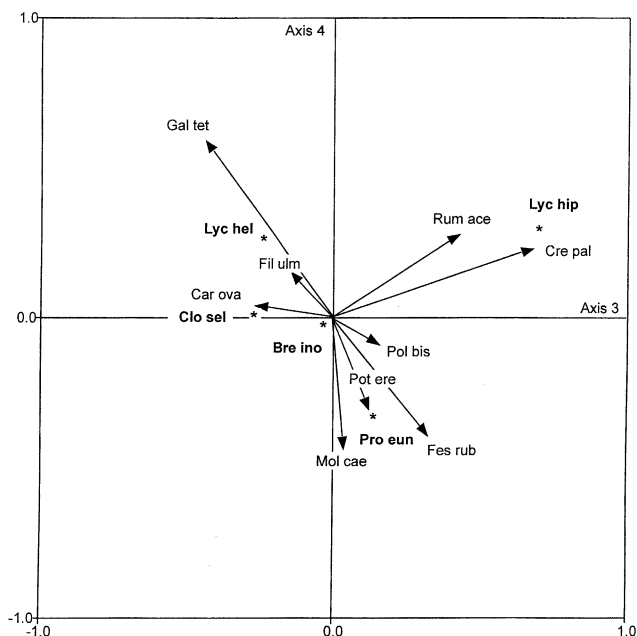


Fig. 3. CCA ordination biplot of the butterfly and plant species on axes 3 and 4. Butterfly and plant species names are abbreviated. Butterfly species are indicated in bold face.

*pothoe* and *P. eunomia*, so simple linear models were selected. All the models were significant at the 0.01 level and explained a considerable fraction of variation.

Table 5 summarises the correlation coefficients between observed abundances and those predicted by the models on the test set. The model significantly explained the observed abundances of four butterfly species. However, the model did not explain *L. hippothoe* abundance well.

Table 3  
Partition of the variation of the butterfly assemblage

Source of variation	Fraction of variation of the butterfly species data
Nonspatial environmental	52%
Spatially structured environmental	24%
Nonenvironmental spatial	2%
Unexplained	22%

Table 4  
Regression equation predicting butterfly abundances from site scores on CCA axes (and on its quadratic expressions) and  $R^2$  showing the proportion of abundance variation that is explained. Log-transformation of abundances increases strongly the  $R^2$  only for *B. ino* and *L. hippothoe*. The parameter ranking corresponds to the entry step during the stepwise multiple regression process

Species	Regression equation	$R^2$
<i>B. ino</i>	$\text{Log}(N) = -0.661 \times \text{CCA1} - 0.560 \times \text{CCA2}^2 - 0.231 \times \text{CCA1}^2 + 3.767$	75%
<i>C. selene</i>	$N = 5.415 \times \text{CCA1} + 2.783 \times \text{CCA1}^2 + 2.734$	86%
<i>L. helle</i>	$N = 4.763 \times \text{CCA2}^2 + 4.013 \times \text{CCA4} - 3.230 \times \text{CCA1} + 5.421$	61%
<i>L. hippothoe</i>	$\text{Log}(N) = 0.546 \times \text{CCA1} + 0.420 \times \text{CCA3} - 0.231 \times \text{CCA2} + 0.888$	62%
<i>P. eunomia</i>	$N = 5.192 \times \text{CCA2} - 4.058 \times \text{CCA1} + 8.270$	63%

Table 5  
Pearson correlation coefficients between observed and predicted abundance, and associated probabilities computed on the test site set

<i>B. ino</i>	<i>C. selene</i>	<i>L. helle</i>	<i>L. hippothoe</i>	<i>P. eunomia</i>
$r$ 0.559	0.642	0.935	0.276	0.729
$P$ 0.024	0.007	0.002	0.179	0.013

The difference between observed and predicted abundances was significantly correlated with site isolation for *L. helle* ( $r^2 = 83\%$ ;  $P = 0.01$ ) and *P. eunomia* ( $r^2 = 79\%$ ;  $P = 0.01$ ). Fig. 4 indicates a similar pattern for these two species. Species abundance predicted by the models was overestimated (predicted abundance greater than observed) in the most isolated sites. In less isolated sites, the predicted abundance was underestimated.

## 4. Discussion

### 4.1. The choice of botanical relevé

Plant communities constitute a promising variable for the description of the habitat quality for butterflies. This statement may be particularly correct for specialist species, intimately linked to certain resources or environmental conditions, with restricted distribution. It remains to evaluate the potential of this kind of models for more widespread and generalist species. Detailed vegetation relevés integrate many potential ecological factors (Ellenberg et al., 1992; Wheeler and Proctor, 2000) governing butterfly species distribution (Oostermeijer and van Swaay, 1998). Some plant species represent essential resources providing nutrients or microclimates, while other plants can also indicate peculiar

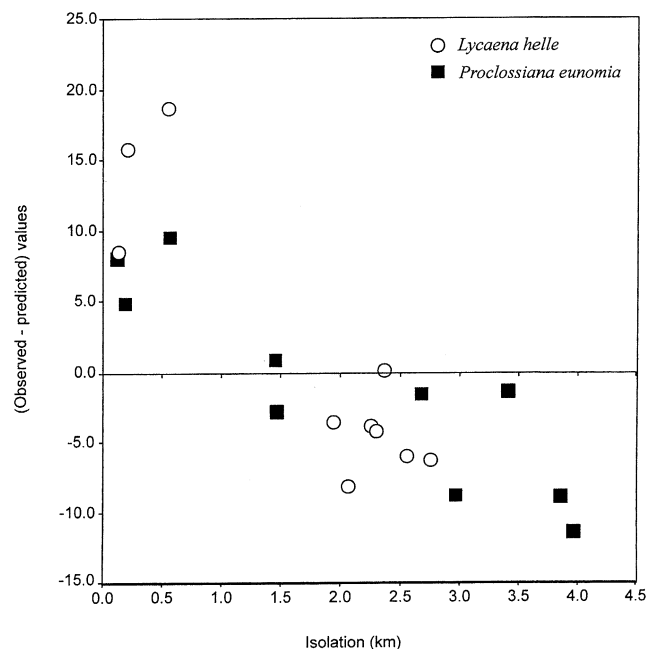


Fig. 4. Model residuals (observed-predicted abundance) vs. isolation (average distance to five closest populations) for *L. helle* and *P. eunomia* in the test site set.

ecological situations that are favourable to butterfly species, even if there is no causal relationship between their simultaneous presence. The host plants are well widespread in the studied area and they alone explain 48% of the butterfly species variation. However, including other plant species as explaining variables enhances the predictive power of the models, which succeeds in explaining 76% of the butterfly variation.

Other ecological variables such as edaphic descriptors could be added to model butterfly distribution. Nevertheless, by considering plant associations as explaining variables, we probably incorporated much information shared with soil characteristics. Plant species are strongly correlated to soil conditions and some plant species can be used as indicators of particular combinations of ecological characteristics (Ellenberg et al., 1992). Plant relevés are preferred because they are more easily obtained from field surveys than detailed analyses of soil composition. On methodological grounds, the use of plant relevés limits spatial heterogeneity at a microlocal scale of soil component analyses and integrates seasonal variations of soil conditions. They also incorporate constraints on land use and management, to which butterflies are also very sensitive. Moreover, a large part of the variation in the butterfly distribution explained by vegetation cannot be explained as spatially structured variation, which emphasises the importance of local conditions. Considerable spatial autocorrelation was demonstrated only for *L. hippothoe*, probably due to its distribution quite well limited to the highest altitude sites in the sampling area. This distribution is strongly correlated to peculiar ecological conditions and it could not be suspected that this pattern was only explained by spatial autocorrelation. In this way, the conditions at a local level are summarised by the composition of the flora that in turn gives an indication of the potential presence of the butterflies.

Our results clearly demonstrate the potential of botanical relevés to evaluate the site quality for butterfly species. Plant frequencies are better predictors of butterfly abundance than soil variables.

#### 4.2. Methodological assessments

In the present work, we developed a model based on CCA and regression analysis to predict species abundances (ter Braak, 1987; Jongman et al., 1995). We chose CCA to describe the distribution of the butterfly assemblage assuming that the five species react to the same environmental gradients, represented by the CCA axes. These composite gradients of vegetation spatial turnover probably express underlying gradients of nutrient content, humidity and pH (Oostermeijer and van Swaay, 1998; Wheeler and Proctor, 2000). The occurrence of the five butterfly species is almost strictly confined to wet meadows. Indeed, the plant species that nourish caterpillars and adults are closely linked to these habitats and widely distributed among them. For example, *P. bistorta*, host plant and nectar source for *L. helle*, *L. hippothoe* and *P. eunomia*, was present at 85% of the sites.

*Filipendula ulmaria*, host plant of *B. ino*, was present at 71% of the sites while the occurrence of *V. palustris*, host plant of *C. selene*, was recorded at 53% of the sites. Furthermore, previous studies showed that butterfly abundances may be roughly explained by two major gradients: one related to pH and concentration of metallic cations, and the other to water table depth (Sawchik, in preparation). A less pronounced fertility gradient was also identified. Complex vegetation gradients probably integrate the effects of these environmental gradients. Thus, we assumed that the butterfly species respond to same environmental gradients at local scale.

We conclude that canonical ordination methods like CCA, combined with regression techniques, are an important tool to model species-habitat associations. This synecological approach allows the selection of a unique and reduced set of predictor variables for a given set of species. This feature makes this approach an interesting option because of the small number of predictor variables to be measured in the field. However, model predictions and implementations must be taken carefully. Model validation is a fundamental phase to evaluate the usefulness, accuracy and generality of models and it needs to be explored with more detail to define model limitations. Nevertheless, the development of this type of models represents an important step to understand the causes that govern the distribution of the species.

#### 4.3. Biological assessments

The five butterfly species are distributed along a gradient from lowland alluvial grassland to high altitude poor grasslands and mires. Ecological niche amplitudes of four of the five species were well described by sampling only wet habitats. Only the *L. hippothoe* ecological niche was probably not completely well explained by botanical relevés. This species showed a nonsignificant correlation between predicted and observed values in the test set. One probable reason is its very low number of suitable sites. Field observations revealed that *L. hippothoe* seems to be also present in neighbouring drier grasslands with many flowering plants. In our study, field monitoring did not completely cover its whole habitat amplitude in the Plateau des Tailles. Finally, this species exhibits such a territorial behaviour that when the site area is lower than 1 ha, often only one pair is regularly observed at each field visit, even in high-quality sites. The other four species were not observed elsewhere and their ecological niche seemed to be well sampled.

The 10 plant species retained in the analysis probably are good indicators of the ecological niche of butterfly species. Species other than host plants probably describe ecological conditions necessary to increase or to constrain butterfly abundance. Some butterfly species need rather specific habitat characteristics to complete their life cycles. For example, Fig. 3 suggests a pattern to explain the distributions of *L. helle* and *P. eunomia*, respectively. The importance of the host plant (*P. bistorta*) for the two butterfly species was clearly stated. In addition, *P. eunomia* abundance was also highly associated with *F. rubra* and *M. coerulea*, while

*L. helle* was associated to *G. tetrahit*. For *P. eunomia*, detailed field observations have demonstrated that *F. rubra*, *M. coerulea* as well as *D. cespitosa* tufts play a crucial role in its life cycle, allowing caterpillars to spend the bad season sheltered and to escape from the early spring floods (Goffart and Waeyenbergh, 1994). For *L. helle*, the presence of *G. tetrahit* denotes a succession from wet meadows to bush and first forest stages. Field observations revealed that *L. helle* was indeed always more frequent in sites protected from dominant winds, by deciduous or spruce trees or bushes in smaller sites, or near the forested borders of greater ones (Goffart, personal communication).

#### 4.4. Evaluation of the isolation effect

Isolation and fragmentation are well known processes acting on local and regional distribution of species with discrete interacting populations (Nève, 1996; Hanski, 1998). Probably, *B. ino*, *C. selene*, *L. helle* and *P. eunomia* are examples of the patchy metapopulation model (Baguette and Nève, 1994; Nève et al., 1996; Hanski and Gilpin, 1997). These four butterfly species occupied almost all the available suitable habitat patches; only very isolated habitat patches seem being empty. For *L. helle* and *P. eunomia*, important isolation effects were demonstrated on the test set. All the sites where the model predicts more specimens than observed, had an average distance to the five nearest occupied sites greater than 2 or 3 km. These isolated sites may show greater chances of being empty indicating that if local populations become extinct, they have low probabilities of being recolonised. This was confirmed for *P. eunomia* where MRR results (Nève et al., 1996) indicated that migration between close patches (< 1 km) was much more frequent than migration between distant patches (> 3 km).

Reciprocally, the less isolated sites showed greater abundances than expected. The higher abundance observed in these highly connected sites may be simply explained by rescue effects (Moilanen, 1998), i.e. the proportion of individuals finding another suitable site is higher when the closer the sites are to each other. The highest geographical concentration of potential sites captures the emigrating individuals more efficiently and reduces dispersal mortality, resulting in higher local abundances.

These results indicate that botanical relevés are good predictors but not sufficient to estimate the abundance of butterfly species in a site. We still need butterfly samplings to assess the effect of the site isolation or to evaluate the significance of a site in an ecological network. The botanical relevé mainly reveals the necessary complementarity between plant and butterfly species needed to evaluate the biological interest of a peculiar site, in a functional ecological network perspective. Landscape structure may have a strong influence on metapopulation viability. Thus, habitat-association models must be evaluated in a landscape context (Hanski and Gilpin, 1997).

#### Acknowledgements

We thank J. De Sloover and J.L. Gathoye for the phytosociological works, A. Peeters (Unité d'Ecologie des Prairies, AGRO/UCL, Belgique) for soil analysis and C. Arsenault, C. Gilot and L. Wargé for the butterfly samplings. P. Goffart, G. Nève, M. Waeyenbergh, D. Calliari, A. Brazeiro, L. Gimenez, E. Cristina, N. Priggioni provided valuable inputs for all the study process and on the first drafts of the manuscript. We thank C. Thomas and two anonymous referees for useful comments on the manuscript. Financial supports were provided by a Uruguayan subvention from CONYCIT and a contract of the Walloon Environment Minister to JS. MD is a beneficiary of a grant and a two year contract as senior research assistant from the National Fund of Scientific Research (Belgium). This is contribution BRCO11 from the Biodiversity Research Center (UCL).

#### References

- Baguette, M., Nève, G., 1994. Adult movements between populations in the specialist butterfly *Proclissiana eunomia* (Lepidoptera, Nymphalidae). *Ecol. Entomol.* 19, 1–5.
- Borcard, D., Legendre, P., Drapeau, P., 1992. Partialling out the spatial component of ecological variation. *Ecology* 73, 1045–1055.
- Boyce, M.S., 1992. Population viability analysis. *Annu. Rev. Ecol. Syst.* 23, 481–506.
- ter Braak, C.F.J., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67, 1167–1179.
- ter Braak, C.F.J., 1987. The analysis of vegetation-environment relationship by canonical correspondence analysis. *Vegetatio* 69, 69–77.
- ter Braak, C.F.J., 1989. Canoco—a Fortran program for canonical community ordination by (partial) (detrended) (canonical) correspondence analysis, principal components analysis and redundancy analysis (version 2.1). Agricultural Mathematics Group, Ministry of Agriculture and Fisheries, Wageningen.
- ter Braak, C.F.J., 1990. Update notes: Canoco version 3.10. Agricultural Mathematics Group, Ministry of Agriculture and Fisheries, Wageningen.
- ter Braak, C.J.F., Prentice, I.C., 1988. A theory of gradient analysis. *Adv. Ecol. Res.* 18, 271–317.
- ter Braak, C.J.F., Smilauer, P., 1998. Canoco reference manual and user's guide to Canoco for Windows: software for canonical community ordination (version 4). Microcomputer Power, Ithaca, NY.
- Burgman, M.A., Ferson, S., Akçakaya, H.R., 1996. Risk Assessment in Conservation Biology. Chapman & Hall, London.
- Chew, F.S., 1980. Food plants preferences of *Pieris* caterpillars (Lepidoptera). *Oecologia* 46, 347–353.
- Daiyuan, P., André, B., Legendre, P., Gérard, D., 1998. Influence of edaphic factors on the spatial structure of inland halophytic communities: a case study in China. *J. Veg. Sci.* 9, 797–804.
- Dowdy, S., Wearden, S., 1985. Statistics for Research. John Wiley & Sons, New York.
- EEW, 1993. Etat de l'environnement wallon, Chapitre Faune-Flore, Ministère de la Région wallonne, Jambes.
- Ellenberg, H., Weber, H.E., Düll, R., 1992. Indicator values of plants in Central Europe, *Scripta Geobotanica*. Goltze Verlag, Göttingen.
- Fischer, K., Beinlich, B., Platcher, H., 1999. Population structure, mobility and habitat preferences of the violet copper *Lycaena helle* (Lepidoptera: Lycaenidae) in Western Germany: implications for conservation. *J. Insect. Conserv.* 3, 43–52.
- Gilbert, L., Singer, M.C., 1975. Butterfly ecology. *Annu. Rev. Ecol. Syst.* 6, 365–397.



- Goffart, P., Waeyenbergh, M., 1994. Exigences écologiques et gestion des populations de deux papillons des prairies humides ardennaises: le cuivré et le nacré de la bistorte (*Lycaena helle*, *Proclissiana eunomia*). Cah. Réserves Nat. 7, 21–29.
- Goffart, P., Baguette, M., de Bast, B., 1992. La situation des Lépidoptères Rhopalocères en Wallonie ou que sont nos papillons devenus? Bull. Soc. R. Belge Entomol. 128, 355–392.
- Hanski, I.A., 1998. Metapopulation ecology. Oxford University Press, Oxford.
- Hanski, I.A., Gilpin, M.E. (Eds.), 1997. Metapopulation Biology: Ecology, Genetics and Evolution. Academic Press, San Diego.
- Holl, K.D., 1995. Nectar resources and their influence on butterfly communities on reclaimed coal surface mine. Restoration Ecol. 3, 76–85.
- Jongman, R.H.G., ter Braak, C.J.F., van Tongeren, O.F.R. (Eds.), 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press, Cambridge.
- Kent, M., Coker, P., 1992. Vegetation Description: A Practical Approach. Belhaven Press, London.
- Kremen, C., 1992. Assessing the indicator properties of species assemblages for natural areas monitoring. Ecol. Appl. 2, 203–217.
- Legendre, P., 1993. Spatial autocorrelation: trouble or a new paradigm?. Ecology 74, 1659–1673.
- Legendre, P., Fortin, M.J., 1989. Spatial patterns and ecological analysis. Vegetatio 80, 107–178.
- Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier, Amsterdam.
- Legendre, P., Vaudor, A., 1991. Le logiciel R: analyse multidimensionnelle, analyse spatiale. Université de Montréal, Montréal.
- Lindenmayer, D.B., Burgman, M.A., Akçakaya, H.R., Lacy, R.C., Possingham, H.P., 1995. A review of the generic computer-programs Alex, Ramas/space and Vortex for modeling the viability of wildlife metapopulations. Ecol. Modell. 82, 161–174.
- Loertscher, M., Erhardt, A., Zettel, J., 1995. Microdistribution of butterflies in a mosaic-like habitat: the role of nectar sources. Ecography 18, 15–26.
- Moilanen, A., 1998. Modeling metapopulation dynamics, Ph.D. Thesis, University of Helsinki, Helsinki.
- Moilanen, A., Hanski, I., 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. Ecology 79, 2503–2515.
- Murphy, D.D., 1983. Nectar sources as constraints on the distribution of egg masses by the checkerspot butterfly. *Euphydryas chalcedona* (Lepidoptera: Nymphalidae). Environ. Entomol. 12, 463–466.
- Nève, G., 1996. Dispersion chez une espèce à habitat fragmenté: *Proclissiana eunomia* (Lepidoptera, Nymphalidae), Ph.D. Thesis, Université catholique de Louvain, Louvain-La-Neuve.
- Nève, G., Barascud, B., Hughes, R., Aubert, J., Descimon, H., Lebrun, P., Baguette, M., 1996. Dispersal, colonisation power and metapopulation structure in the vulnerable butterfly *Proclissiana eunomia* (Lepidoptera: Nymphalidae). J. Appl. Ecol. 33, 14–22.
- Øksanen, J., Minchin, P.R., 1997. Instability of ordination axes results under changes in input order: explanations and remedies. J. Veg. Sci. 8, 447–454.
- Oostermeijer, J.G.B., van Swaay, C.A.M., 1998. The relationship between butterflies and environmental indicator values: a tool for conservation in a changing landscape. Biol. Conserv. 86, 271–280.
- Palmer, M.W., 1993. Putting things in even better order: the advantages of canonical correspondence analysis. Ecology 74, 2215–2230.
- Pollard, E., 1977. A method of assessing changes in the abundance of butterflies. Biol. Conserv. 12, 115–134.
- Porter, K., 1992. Eggs and egg-laying. In: Dennis, R.L.H. (Ed.), The Ecology of Butterflies in Britain. Oxford University Press, Oxford, pp. 46–72.
- Post, E., Stenseth, N.C., 1998. Large-scale climatic fluctuation and population dynamics of moose and white-tailed deer. J. Anim. Ecol. 67, 537–543.
- Potvin, C., Roff, D.A., 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics. Ecology 74, 1617–1628.
- Ravenscroft, N.O.M., 1994. The ecology of the chequered skipper butterfly *Carterocephalus palaemon* in Scotland: food plant quality and population range. J. Appl. Ecol. 31, 623–630.
- SAS, 1982. SAS users guide: statistics, SAS. Institute Inc., Cary, North Carolina.
- Sawchik, J., 1999. Répartition spatiale et dynamique de populations de rhopalocères des prairies humides du Plateau des Tailles, Ph.D. Thesis, Université catholique de Louvain, Louvain-la-Neuve.
- Shaffer, M.L., 1990. Minimum viable population: coping with uncertainty. In: Soulé, M.E. (Ed.), Viable Populations for Conservation. Cambridge University Press, Cambridge, pp. 69–86.
- Shreeve, T.G., 1992. Adult behaviour. In: Dennis, R.L.H. (Ed.), The Ecology of Butterflies in Britain. Oxford University Press, Oxford, pp. 22–45.
- Smallidge, P.J., Leopold, D.J., Allen, C.M., 1996. Community characteristics and vegetation management of karner blue butterfly (*Lycaeides melissa samuelis*) habitats on rights-of-way in east-central New York, USA. J. Appl. Ecol. 33, 1405–1419.
- Thomas, J.A., 1984. A quick method for estimating butterfly numbers during surveys. Biol. Conserv. 27, 195–211.
- Thomas, J.A., Bourn, N.A.D., Clarke, R.T., Stewart, K.E., Simcox, D.J., Pearman, G.S., Curtis, R., Goodger, B., 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proc. R. Soc. Lond. B 268, 1791–1796.
- Wheeler, B.D., Proctor, M.C.F., 2000. Ecological gradients, subdivisions and terminology of north-west European mires. J. Ecol. 88, 187–203.