

## Towards indicators of butterfly biodiversity based on a multiscale landscape description

J.-P. Rossi<sup>\*</sup>, I. van Halder

INRA, UMR1202 BIOGECO, F-33612 Cestas, France

### ARTICLE INFO

#### Article history:

Received 12 January 2009

Received in revised form 26 July 2009

Accepted 27 July 2009

#### Keywords:

Butterfly species richness

Biodiversity

Landscape metrics

Spatial scales

Partial Least Squares Regression

### ABSTRACT

The conservation of biodiversity has gained prominence in ecological research for the last decades. Conservation actions require a measure of biodiversity such as species richness, but its assessment is very difficult, even for small areas and therefore the search for surrogates (i.e. indicators) of biodiversity has emerged as an active research topic. We investigated the relationships between butterfly species richness and landscape structure and composition in two pine plantation sites in Southwest France. We assessed the correlation between butterfly species richness and a set of 15 landscape metrics computed for 18 land-uses at 10 different spatial scales. Spatial scales were accounted for by computing landscape metrics for circular buffers with radius ranging from 100 to 1000 m. The joint use of the Partial Least Squares Regression (PLSR) and a stepwise regression procedure revealed strong correlations between butterfly species richness and various landscape metrics in both study sites. The selected landscape metrics differed from one site to another and mostly involved measures of landscape fragmentation. We found a very strong effect of the spatial scale of investigation upon the perception of the landscape–butterfly richness relationship. Our main conclusions are that (i) certain landscape attributes can potentially serve as indicators for butterfly species richness at the landscape scale; (ii) future indicators of biodiversity based on landscape features should consider various spatial scales.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

The conservation of biodiversity has gained prominence in ecological research for the last decades and it is widely recognized that large-scale habitat alteration has led to unprecedented worldwide species extinction (Pimm et al., 1995). Conservation actions require a measure of biodiversity such as species richness, but its assessment is very difficult, even for small areas. As a consequence, the search for surrogates (i.e. indicators) of biodiversity has emerged as an active research topic (Araújo et al., 2001; Prendergast and Eversham, 1997). Biodiversity indicators can be based on species richness of one or several taxonomic groups, with the assumption that the diversity of this group is correlated with that of other groups (Maes and Van Dyck, 2005). The level of correlation depends on the taxonomic resolution employed as well as the taxa themselves and the results have been contrasted so far (Similä et al., 2006). The literature contains case studies where some indicator groups have been identified (Dynesius and Zinko, 2006; Schmidt et al., 2006) but other authors reported weak correlations (Chase et al., 2000; Wolters et al., 2006) or highly spatially variable and hence unpredictable relationships (Prendergast and Eversham, 1997). A

different approach consists in using habitat patch characteristics and landscape features as an indicator of species richness (Dauber et al., 2003). The basic idea is that since landscape acts upon biodiversity dynamics there must be some landscape features that can be correlated to biodiversity and eventually be used as indicators (Araújo et al., 2001).

The biodiversity at the local scale (e.g. the scale of the habitat patch) depends primarily upon local habitat characteristics but also on communities scattered across the landscape. This corresponds to the idea that local biodiversity does not only depend on local processes (e.g. local extinction) but also on processes that act over regional or biogeographical scales (e.g. migration) (Ricklefs, 1987). The effects of landscape configuration (i.e. the “matrix effect”) correspond to a variety of processes among which dispersal, source–sink dynamics, neighborhood effects and metapopulation dynamics (Dunning et al., 1992). Fragmentation (e.g. With and King, 1999b) and matrix quality (e.g. Fahrig, 2001) can strongly affect species extinction thresholds while lacunarity alters colonization success hence the dynamics of biodiversity (With and King, 1999a). At the regional scale, the so-called  $\gamma$  diversity is strongly dependent on landscape composition (Dunning et al., 1992) with subsequent effects at local scales (Whittaker, 1972). The impacts of landscape are therefore multiple, and depend on the taxa examined, the scale considered and the scale at which organisms interact with their environment (Bestelmeyer et al., 2003).

<sup>\*</sup> Corresponding author.

E-mail address: [Jean-Pierre.Rossi@pierroton.inra.fr](mailto:Jean-Pierre.Rossi@pierroton.inra.fr) (J.-P. Rossi).

Butterflies are insects that form communities including species with different habitat preferences and dispersal capacities and therefore contrasted response to landscape composition and configuration (Dennis et al., 2003). Most butterfly species need complementary resources like host plants (for larvae), nectar plants (for adults) and sites for resting or overwintering. In addition, many species use resources in different vegetation types and/or are mostly active along edges. As a consequence, there are direct and strong relationships between butterfly community structure and richness and habitat characteristics, vegetation composition and landscape (Dennis et al., 2003). Butterflies are therefore a very convenient group to study the relationships between landscape mosaic and species richness (see e.g. Debinski et al., 2001; Schneider and Fry, 2001) and explore the potentials for biodiversity indicators.

Building a landscape-based indicator of butterfly biodiversity amounts to (i) describing landscape by means of relevant metrics, (ii) correlating the metrics with a biodiversity measure such as species richness. There is no way to determine a priori which scales matter to the taxa at hand (only a range of known or likely spatial scales can be suggested) and one metric may be linked to species richness at one scale while another metric would be at a different scale, therefore a large range of scales must be investigated. Another difficulty arises when choosing the landscape metrics that must be computed and used in further correlation analyses: their inner pair correlations can cause problems in the data processing (i.e. collinearity between landscape metrics) (Schindler et al., 2008; Cushman et al., 2008). Moreover, the multiscale behavior of landscape metrics themselves implies multiscale analysis to adequately characterize and monitor landscape heterogeneity (Wu, 2004). As an example, Zaccarelli et al. (2008) emphasized the insights gained from a multiscale approach in the establishment of regional conservation network to manage biodiversity at the regional scale. Both theoretical and empirical studies show that indicators of species richness based on landscape metrics should be multiscale although there are only few case studies (e.g. Bergman et al., 2004; Krauss et al., 2003) and some questions regarding the statistical methodology needed to tackle these issues remain to be investigated.

The aim of the present study was to explore the relationships between butterfly species richness and landscape metrics measured at different spatial scales. We based our analyses on butterfly communities sampled in two sites located in the same region (Southwestern France) where the landscape is dominated by pine plantation forests with virtually no variability in topography (Van Halder et al., 2008). Our survey is a first step towards more general indicators that require more replicated studies and additional sampling in various landscape types. In this paper we explore the contribution of a multiscale view of the landscape–biodiversity relationship, concentrate on the statistical issues and introduce the use of the Partial Least Squares Regression (PLSR) (Wold et al., 2001) as a tool to design landscape-based indicators.

## 2. Materials and methods

The present study is based on data published by Van Halder et al. (2008). Study sites and sampling procedures are fully described in the original paper and therefore we only provide the main information here.

### 2.1. Landscapes

The study was carried out in Southwestern France in the “Landes de Gascogne” (Van Halder et al., 2008). The region covers ca. 1 million ha of plantation forest of maritime pine (*Pinus pinaster*, Aiton 1789). Landscapes are therefore dominated by pine

plantations with rare deciduous woodlands mostly found along rivers (riparian forest) or as scattered patches of a few hectares. Open areas comprise maize fields, pine clear cuts and firebreaks. Soils are acid podzols (pH of 3.5–5.5). Pine forests are submitted to intensive silvicultural managements including fertilization, mechanical understorey removal and four thinning operations during the rotation cycle (Van Halder et al., 2008).

### 2.2. Study sites

We investigated two sites (namely the “Tagon” and the “Solferino”) located South West of Bordeaux and covering 5000 ha and 10,500 ha, respectively. Soils were similar but landscape composition and structure differed to a certain extent. Butterflies were studied in 33 and 46 sampling plots in Solferino and Tagon, respectively. Each plot corresponded to one of the seven main habitat types (land-uses) commonly encountered in the Landes de Gascogne (Van Halder et al., 2008). Five habitat types depicted the main stages of the maritime pine plantations in Southwestern France: herbaceous and shrubby clearcuts and three size classes of pines (young pines, canopy height  $\leq 7$  m; mid-class pines, canopy height 7–15 m; and older pines, canopy height  $> 15$  m). Other habitats included deciduous woodlands (isolated patches or riparian forests) and firebreaks or powerlines (hereafter called firebreaks).

### 2.3. Sampling procedure

Butterflies were sampled using a line-transect protocol that is fully described in Van Halder et al. (2008). Butterfly sampling was carried out within and at the edge of each sampling plot using two 400 m long transects. Butterflies were counted within 2.5 m on each side of the transect line and 5 m ahead of the recorder. Each plot was visited four times between May 14th and September 4th 2004.

### 2.4. Landscape description

Land-use types in the two study sites were mapped in a GIS based on aerial colour photos with a resolution of 50 cm as background layer. Photos dated from 2000 and 2002 for Tagon and Solferino, respectively. Field surveys were conducted to check patch attributes and correct for changes that might have occurred since the date of the photos. Aerial photos were mapped using 18 land-use categories including the 7 main habitat types.

### 2.5. Data analysis

#### 2.5.1. Basic statistics

If both study sites hosted similar butterfly communities it would imply that all data could be analyzed as a whole. On the contrary, different communities correspond to different “sample populations” and as such deserve separate landscape–species richness analyses. We used a correspondence analysis (CoA) (Legendre and Legendre, 1998) to determine to which extent the community structure differed amongst the sites of Solferino and Tagon. The between-sites difference was assessed using an inertia test based on 999 randomizations (Manly, 1997). This was done using the software `R` (R Development Core Team, 2008) and the `R` package `ade4` (Chessel et al., 2004). We used the bootstrap to re-estimate the species richness and approximate the bias in the estimator of species richness (Manly, 1997, p. 36). This allowed to compute bias-corrected species richness and the associated confidence interval (Manly, 1997). We compared the observed species richness of the two sites by means of a randomization test following Manly (1997, p. 6).

**Table 1**

List of the metrics used to describe the plantation forest landscapes in two sites located in Southwestern France. Acronyms, names and codes correspond to the terminology of McGarigal et al. (2002). C and L respectively stands for class and landscape level metrics (see McGarigal et al., 2002, for a complete description of each metric).

Acronym	Name	Level	Code
PLAND	Proportion of landscape	C	C4
PD	Patch density	C, L	C6, L6
LPI	Largest patch index	C	C10
ED	Edge density	C, L	C8, L8
PRD	Patch richness density	L	L125
SHDI	Shannon's diversity index	L	L127
SIDI	Simpson's diversity index	L	L128
SHEI	Shannon's evenness index	L	L130
SIEI	Simpson's evenness index	L	L131
AI	Aggregation Index	L	L116
COHESION	Patch cohesion index	L	L121

### 2.5.2. Multiscale landscape description

Landscapes were described around each sampling point using circular areas referred to as buffers. We used buffers of increasing radii so as to capture landscape features at different spatial scales. Buffers radii ranged from 100 to 1000 m. The resulting micro-landscapes were described using landscape metrics reflecting both structure and composition. The metrics are listed in Table 1 and fully described in McGarigal et al. (2002). Overall, each sampling location was described using 11 landscape metrics and 4 class metrics computed for each of the 18 landscape classes, i.e. the land-use categories (Table 1). Given that we used 10 buffer sizes the resulting landscape data set comprised  $(11 + 4 \times 18) \times 10 = 830$  variables.

### 2.5.3. Relating landscape metrics to species richness

Van Halder et al. (2008) showed that the butterfly species community was strongly affected by habitat in the “Landes de Gascogne” and therefore this effect might obscure the relationship between species richness and surrounding landscape (matrix effect). We accounted for that difficulty by removing the average habitat effect through centering and therefore based our analyses on the habitat-centered species richness (referred to as species richness). Because the number of landscape metrics was larger than the number of objects (i.e. the sampling locations), the classical regression analyses (e.g. multiple linear regression) could not be used. Another characteristic of our data sets is that they are likely to contain many noisy, partially redundant or collinear variables (Cushman et al., 2008). Such data can be analyzed by Partial Least Squares Regression (PLSR). Basically, PLSR is a linear regression of one or more response variables  $\mathbf{y}$  onto a number of principal component scores from a predictor matrix  $\mathbf{X}$  but whereas principal component analysis maximizes the variance of the scores, PLSR maximizes the covariance between the scores and the response (Mevik, 2006). It is a standard tool in chemometrics where large data sets with strong collinearity are common (Wold et al., 2001) and various applications in other fields of life sciences have been published (e.g. analysis in brain images: McIntosh et al., 1996, tumour classification: Nguyen and Roche, 2002). A detailed introduction to the PLSR can be found in Wold et al. (2001), Mevik and Wehrens (2007) and Martens (2001).

We used the PLSR in a design strictly similar to a multiple regression,  $\mathbf{X}$  being the landscape metrics table and  $\mathbf{y}$  the species richness. The computations were done using the software R (R Development Core Team, 2008) and the associated package `pls` (Mevik and Wehrens, 2007). The PLSR leads to a statistical model linking the response variable (species richness) to all the explanatory variables and additional computations are needed in order to identify the pertinent explanatory variables (Gauchi

and Chagnon, 2001). This step is important because it allows to reduce the huge initial data set to a lower number of pertinent variables that can be used in further analysis like standard multiple regressions. We used the Variable Importance in the Projection (VIP) approach described in Chong and Jun (2005) and used the threshold value of  $VIP = 1$  for selecting the variables deemed pertinent. The computations were done using the software R (R Development Core Team, 2008) and the associated function `VIP.R` written by Bjørn-Helge Mevik and available at <http://mevik.net/work/software/VIP.R>. Since the VIP approach does not include a formal statistical test, the metrics of interest were subsequently incorporated into a multiple stepwise regression with the aim to build a simple model linking landscape and species richness.

When a metric had a VIP value greater than one (for Tagon) or two (for Solferino, see Section 3) for different buffer sizes we retained the metric corresponding to the largest VIP value in order to reduce the risk of collinearity.

## 3. Results

### 3.1. Species richness and inter-site variability

The two study sites harbored contrasted species richnesses ranging from 37 (Solferino) to 45 species (Tagon) (Table 2). Both uniques and singletons were low and the between-site difference in species richness remained marked when the bias-corrected estimators were considered. The difference between the observed species richness in Solferino and Tagon was significant ( $p = 0.009$ , randomization test). Both sites hosted butterfly communities that differed in their richness but also in terms of community structure. The CoA showed discrepancies between the sites along axis 2 whereas axis 1 mostly conveyed within sites heterogeneity (Fig. 1). The between-site inertia test performed on CoA indicated that these differences were significant ( $p = 9 \times 10^{-4}$ ). Thirty-three species (67.3% of the total) were common in the two sites. We concluded that the sites of Solferino and Tagon should be analyzed separately given their differences in community structure and species richness.

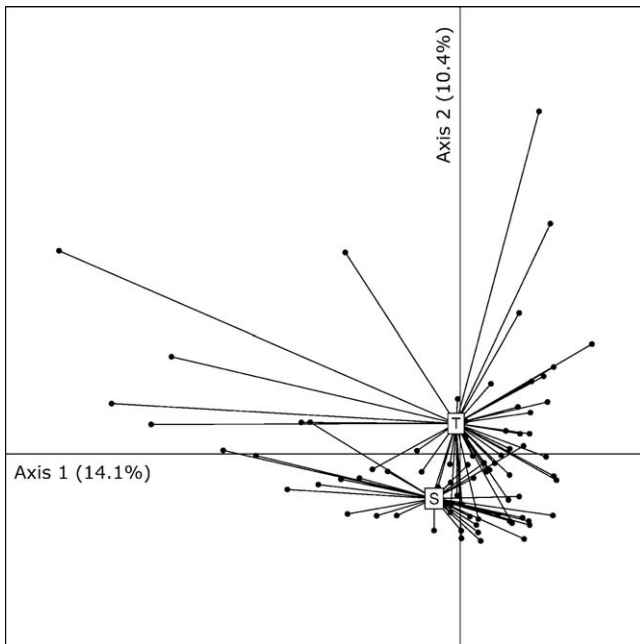
### 3.2. PLSR of the species richness with landscape metrics

The PLSR performed with the species richness as the response variable and the landscape metrics as explanatory variables yielded a first component (i.e. latent variable) that explained 9.6% and 30.9% of the variance of the response variable in the sites of Solferino and Tagon, respectively. The number of landscape metrics with  $VIP \geq 1$  was 131 and 47 (i.e. 15.8% and 5.6% of the 830 initial descriptors available) in Solferino and Tagon. The examination of how the VIP of a given landscape metric changed according to the buffer size (i.e. the scale considered) allowed to explicitly address the question of the effect of the spatial scale upon metric's contribution to the model performance. The impact of the spatial scale was very clear in both Solferino and Tagon as shown in Figs. 2 and 3. These graphs revealed that the VIP of some metrics changed

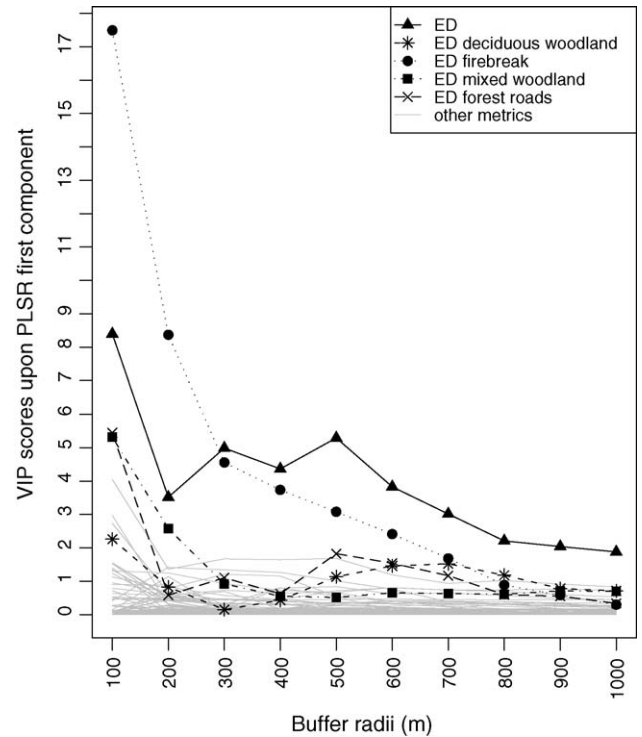
**Table 2**

Observed and estimated butterfly species richness in two sites in the “Landes de Gascogne” (Southwestern France). Singleton and unique are respectively the number of species that were represented by one individual or occurred in one sample only.  $S_b$  is the bias-corrected estimator of the species richness. LB and UB are respectively the upper and lower bounds of the species richness estimator based on 999 randomizations.

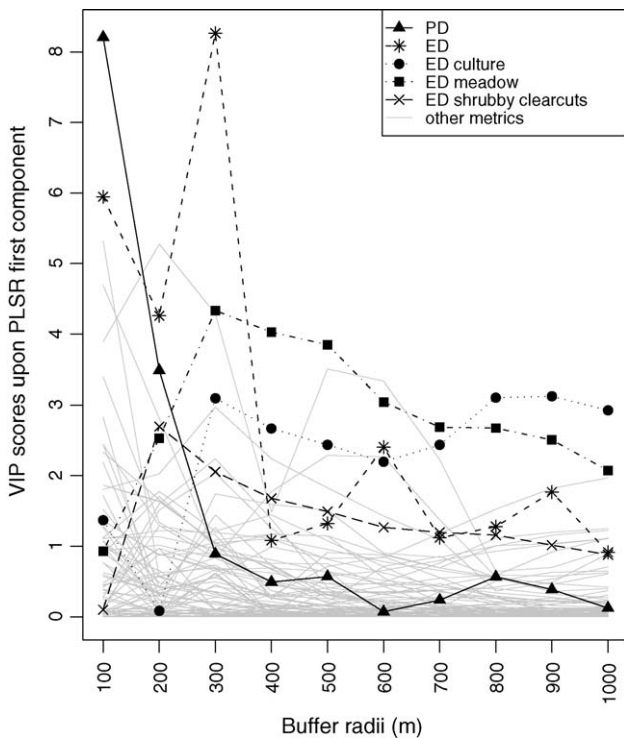
Site	S	$S_b$	LB	UB	Singleton	Unique	n
Solferino	37	39.9	36.9	42.9	3	6	33
Tagon	45	46.2	43.9	48.3	1	1	46
All pooled	49	50.8	48.4	53.2	3	3	79



**Fig. 1.** Correspondence analysis (CoA) of butterfly abundance data in two study sites dominated by pine plantations in Southwestern France. Projection of the sampling units upon the factorial plane 1–2. Open squares are placed at the centre of gravity of the scatter of points corresponding to each site. Lines link samples to the corresponding sites. S: Solferino. T: Tagon.



**Fig. 3.** Output of the PLSR linking butterfly species richness and 830 metrics describing landscape along a range of spatial scales in the site of Tagon (Southwestern France). Relationship between the VIP scores of the landscape descriptors and the spatial scale at which they are computed.



**Fig. 2.** Output of the PLSR linking butterfly species richness and 830 metrics describing landscape along a range of spatial scales in the site of Solferino (Southwestern France). Relationship between the VIP scores of the landscape descriptors and the spatial scale at which they are computed. Changes of the VIP scores of some selected metrics according to the buffer size illustrate scale-dependent fluctuations of the strength of the relationship between landscape descriptors and species richness.

dramatically according to the buffer size considered. In Solferino, for instance, the importance of the edge density (ED) of the shrubby clearcuts was maximum for buffers of 200 m and tended to decrease regularly with increasing radii. The importance of the patch density (PD) decreased sharply with increasing buffer size and the edge density of the meadow was minimum for buffers  $\geq 200$  m, reaching its peak at 300 m and smoothly fluctuated afterwards (Fig. 2). Some landscape descriptors were clearly more strongly related to species richness when measured within small buffers like the edge density of the firebreaks and most of the other metrics measured at the site of Tagon (Fig. 3). The VIP of the edge density of deciduous woodlands decreased for buffer sizes ranging from 100 to 300 m and displayed a bell-shaped curve for larger scales (Fig. 3). Overall, the scale effect manifested itself for a large part of the landscape descriptors and the shape of the curves appeared as very variable according to the metric and the study site considered.

### 3.3. Linear models linking butterfly species richness and landscape metrics

The PLSR and the VIP method led to 131 metrics of interest (i.e.  $VIP \geq 1$ ) in the site of Solferino. Since the number of objects was lower, it was impossible to use the classical regression models and we consequently used a VIP threshold of 2 which led to 17 explanatory variables. Table 3 shows the results of the models fitted for the habitat-centered richness in Solferino and Tagon. The models were significant in both cases and the proportion of explained variance (adjusted  $R^2$ ) ranged from 16% to 41% in Tagon and Solferino, respectively (Table 3). The number of landscape metrics that were finally retained through the stepwise fitting was small and consisted of only one metric (the edge density of firebreaks) in the case of Tagon. In Solferino the species richness



**Table 3**

Linear models linking landscape metrics and butterfly species richness in 2 sites in the “Landes de Gascogne” (Southwestern France). The habitat-centered richness (see text for details) was linearly related to various landscape metrics measured at different spatial scales. Acronym indicates the label of the landscape metrics following McGarigal et al. (2002). C and L respectively stands for class and landscape level metrics. Land-use corresponds to the landscape class name. Buffer radius indicates the radius of the circular area used to compute the landscape metrics.

	Adjusted $R^2$	$p$	Acronym	Level	Land-use	Buffer radius (m)	$p$
Solferino							
Habitat-centered richness	0.407	$2.4 \times 10^{-3}$	ED	L	–	300	$2.2 \times 10^{-2}$
			PD	C	Forest road	100	$3.1 \times 10^{-2}$
			PD	C	Culture	100	$3.5 \times 10^{-2}$
			ED	C	Hedge	300	$3.0 \times 10^{-2}$
Tagon							
Habitat-centered richness	0.1558	$1.8 \times 10^{-2}$	ED	C	Firebreak	100	$2.4 \times 10^{-2}$

was correlated to the patch density of forest roads and of cultures, and edge density of hedges. The scales of importance as represented by the buffer sizes varied from 100 to 300 m according to the landscape metrics considered. The edge density (ED) was the only landscape level metric correlated to species richness (Solferino).

## 4. Discussion

### 4.1. Correlates between butterfly species richness and landscape metrics

The results reported in this work show the existence of a clear statistical correlation between butterfly species richness and certain landscape descriptors i.e. metrics. There is a wealth of literature dealing with the mechanisms at work in the community-landscape relationships (see Dunning et al., 1992, and references therein). Species richness at one place can be affected by the presence and extent of patches of species rich habitat in the surroundings through the so-called “vicinism” or “mass effect” (Zonneveld, 1995). In this situation a flow of individuals originates from species rich habitat patches and allows the presence of species in neighboring habitats where they are not self-maintaining (Shmida and Wilson, 1985). This corresponds to the “spillover effect” by which species can be maintained in unsuitable habitat patches in the framework of source–sink dynamics (Holt, 1997). Other ecological processes affecting communities are landscape complementation and landscape supplementation (Dunning et al., 1992) that occur when individuals move to use non-substitutable or substitutable resources located in different patches. Most butterfly species use complementary resources and therefore the latter processes are likely to shape their communities and affect species richness.

The presence of species in a patch depends on the quality of the patch and on the surrounding landscape (Dauber et al., 2003). In this study we focused on landscape effects by fitting our model on the habitat-centered richness. In the site of Tagon the results clearly emphasized the importance of the firebreaks. The edge density of this habitat in the surroundings of the sampling sites affected the species richness. Interestingly, the multiscale approach adopted here indicated that the near surroundings had the strongest impact with higher correlation for buffers of 100 m radius. Firebreaks constitute rich habitats and as such sampling points corresponding to that land-use harbored the highest biodiversity (see a detailed discussion in Van Halder et al., 2008). In the present study, we showed that this land-use also affected species richness when located in the vicinity of the sampling locations (at least in the close neighborhood i.e. buffers of 100 m). Firebreaks are particularly rich mostly because they are characterized by a more diversified herbaceous vegetation hence

more diverse flowering species (the nectar is an important resource for adult feeding of various butterfly species) as well as more host-plants (important for larval development) (Van Halder et al., 2008). The effect was maximal at small distances (100 m) but remained high for distances up to several hundred of meters (Fig. 3). The species that inhabit firebreaks and contribute to species richness of other, very different habitat types (e.g. deciduous forest patches), may mostly be found at the edges of these patches where local microclimatic conditions are more similar to that of firebreaks. Such a pattern would suggest a typical mass effect and could be properly examined in future analyses.

The data collected at the site of Solferino showed that richness was correlated to landscape edge density. This suggested that a certain level of habitat fragmentation affected butterfly species richness possibly due to increased availability of herbaceous strips along stand edges. Other metrics of significance included patch density of forest roads and cultures as well as edge density of hedges measured within buffers of 100 or 300 m. This results again suggested that open habitats or at least herbaceous strips along stand edges affected butterflies. The main difference between our study sites is that Solferino only featured a very low amount of firebreaks. This may explain the between-site differences in species richness and community structure as well as the discrepancies in the fitted models. On the other hand, it must be noted that our results showed the importance of landscape metrics that measure landscape fragmentation (edge density, patch density) in both study sites.

### 4.2. Scale effects

An important result of our study was the critical effect of the spatial scale at which the landscape is quantified upon our perception of landscape–biodiversity relationships. In the case of butterflies, various studies examined this relationship using different buffer sizes but led to contradictory results (Bergman et al., 2004; Krauss et al., 2003; Weibull et al., 2000). This can be explained by the diversity of ecological processes that tune the local values of species richness and the fact that they take place at various spatial and temporal scales and highly vary in function of local and historical contingencies (Belyea and Lancaster, 1999). As a consequence there is probably no a priori or general spatial scales that could be well adapted to a given taxon. The changes in the level of metrics correlation with biodiversity according to the size of the buffer was dramatic in our study. Previous studies that used different buffer sizes usually compared the performances of each buffer size (e.g. Dauber et al., 2003) whereas we identified the scale at which each metric is the most distinctly linked with species richness. The difference is essential if we consider that the importance of the metrics change dramatically and a priori unpredictably with scale. Figs. 2 and 3 clearly illustrate that there is

no unique (single) buffer size for which all the metrics exhibit their best relationship with species richness. We argue that this is the manifestation of the diversity of the processes that are responsible for community dynamics as well as the multiple scales at which they act. Short-range scale effects may imply processes like mass effect or habitat complementation/supplementation. Relationships between species richness and landscape metrics quantified at larger spatial scales may rather involve other processes like metapopulation dynamics (Holt, 1997). In this study, the correlation between landscape metrics and butterfly richness was in general higher for small buffers which suggested that short-range processes were more important than long-range processes. This may be explained by the fact that our study sites are rather homogeneous (monospecific pine plantation forest) with all land-uses hosting butterflies. This is different from landscapes constituted by alternate patches of habitat and non-habitat where long-range processes like metapopulation dynamics may prevail.

#### 4.3. Statistical issues

An important aspect of landscape analysis lies in that landscape description often involves a large number of partially redundant and strongly interrelated metrics (Cushman et al., 2008). In the type of studies we carried out, explanatory variable redundancy may be troublesome in the multiple stepwise regressions (Graham, 2003). Many problems could be avoided by selecting metrics on the basis of their universality and consistency (Cushman et al., 2008) thereby reducing the number of explanatory variables. However, using a range of buffer sizes necessarily leads to redundancy. We dealt with that problem by using the PLSR, a multivariate tool that is common in chemometrics albeit still rarely used in ecology. The PLSR is not impaired by collinear explanatory variables and as such ideally suits our needs. Another advantage is that PLSR allows to handle data sets that have more descriptors than objects which is likely to occur when numerous metrics are computed with numerous landscape classes. We believe that PLSR is a very promising approach in the field of landscape description because metrics are intrinsically redundant and often constitute very large data sets with more variables than samples. Our results indicate that the outputs of the PLSR can be used to identify those metrics that are the most consistently linked to the response variable and then other, additional data treatments can be done with that reduced set of explanatory variables.

#### 4.4. Landscape as indicator of butterfly species richness

By working on the habitat-centered species richness we accounted for the average habitat type effect but our models could be possibly improved by adding ecological information about habitat quality. This can be done by adding plot scale variables like e.g. understorey vegetation descriptors. Van Halder et al. (2008) showed that a significant part of butterfly community variation was linked to understorey vegetation (see also Collinge et al., 2003, for grassland butterflies). In their study, Dauber et al. (2003) showed the clear prevalence of habitat descriptors upon landscape descriptors and that the latter variables alone performed poorly as predictors of biodiversity. However, our multiscale approach allowed to build simple and significant models which may be improved by adding local descriptors of habitats. The selected landscape metrics differed according to the site considered and this can be explained by differences in the landscape composition (Solferino notably lacked firebreaks). In order to find indicators that are more generally applicable one could search landscape metrics that are related to species richness in both sites. In our landscapes some degree of fragmentation seemed to affect species richness and metrics like patch or edge density may be correlated to species richness in both

sites. More generally, the feasibility of general indicators of biodiversity on the basis of landscape metrics alone remains an open question. The first point is to determine to which extent one or a few taxa can be taken as surrogates for the overall biodiversity. Various studies indicated that such surrogates are rarely found (Chase et al., 2000; Wolters et al., 2006). Our results showed that if such a group was identified, a multiscale approach could reveal the best landscape predictors. Since landscape metrics are always redundant to a large extent one may consider dropping a certain metric and adding another one, with a weaker relationship with the response variable but more easily measurable, universal or consistent (Cushman et al., 2008) for the sake of feasibility. It must be noted that our results and conclusions are highly dependent on the classification scheme used to elaborate the landscape maps. In other words, the definition of the land-use types have an extremely strong influence on the result of the landscape analysis (Turner et al., 2001) hence upon our perception of the landscape–butterfly richness relationships. Because inter-taxa species richness correlation is generally low (Wolters et al., 2006), future indicators may integrate different taxonomic groups, different habitat descriptors and allow for a multiscale landscape matrix description. Another point to consider in future research is that focusing on the total species richness may not be the best strategy. Species richness encompasses very different elements amongst which very common, rare or endangered species. These taxa obviously do not have the same value for conservation purposes which is often why indicators of biodiversity are needed (Pearman et al., 2006). Locally rare species can be very numerous (Novotný and Basset, 2000; Rossi et al., 2006) and their presence is likely to be partly explained by mass effects (Novotný and Basset, 2000). Therefore, using a modified value of species richness including a correction term weighting for the presence of rare species (Colwell and Coddington, 1994) may be a promising approach.

## 5. Conclusions

The present study revealed strong effects of the spatial scale at which landscape metrics are computed upon the outputs of correlative analyses aiming at relating butterfly species richness and landscape composition and structure. The consequences of these results are particularly important when searching for landscape metrics correlated to biodiversity with the aim of elaborating biodiversity indicators.

## Acknowledgements

We thank Dr. C. Kerdelhué and two anonymous reviewers who provided very helpful comments on the manuscript. This study was funded by the GIP ECOFOR (programme “Biodiversité et gestion forestière – Indicateurs de biodiversité forestière”).

## References

- Araújo, M.B., Humphries, C.J., Densham, P.J., Lampinen, R., Hagemeyer, W.J.M., Mitchell-Jones, A.J., Gasc, J.P., 2001. Would environmental diversity be a good surrogate for species diversity? *Ecography* 24, 103–110.
- Belyea, L.R., Lancaster, J., 1999. Assembly rules within a contingent ecology. *Oikos* 86, 402–416.
- Bergman, K.O., Askling, J., Ekberg, O., Ignell, H., Wahlman, H., Milberg, P., 2004. Landscape effects on butterfly assemblages in an agricultural region. *Ecography* 27, 619–628.
- Bestelmeyer, B.T., Miller, J.R., Wiens, J.A., 2003. Applying species diversity theory to land management. *Ecol. Appl.* 13, 1750–1761.
- Chase, M.K., Kristan, W.B., Lynam, A.J., Price, M.V., Rotenberry, J.T., 2000. Single species as indicators of species richness and composition in California coastal sage scrub birds and small mammals. *Conserv. Biol.* 14, 474–487.
- Chessel, D., Dufour, A.-B., Thioulouse, J., 2004. The ade4 package-I: one-table methods. *R News* 4, 5–10.
- Chong, I.-G., Jun, C.-H., 2005. Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* 78, 103–112.

- Collinge, S.K., Prudic, K.L., Oliver, J.C., 2003. Effects of local habitat characteristics and landscape context on grassland butterfly diversity. *Conserv. Biol.* 17, 178–187.
- Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. *Philos. Trans. R. Soc. Lond. B* 345, 101–118.
- Cushman, S.A., McGarigal, K., Neel, M.C., 2008. Parsimony in landscape metrics: strength, universality, and consistency. *Ecol. Indicators* 8, 691–703.
- Dauber, J., Hirsch, M., Simmering, D., Waldhardt, R., Otte, A., Wolters, V., 2003. Landscape structure as an indicator of biodiversity: matrix effects on species richness. *Agric. Ecosyst. Environ.* 98, 321–329.
- Debinski, D.M., Ray, C., Saveriaid, E.H., 2001. Species diversity and the scale of the landscape mosaic: do scales of movement and patch size affect diversity? *Biol. Conserv.* 98, 179–190.
- Dennis, R.L.H., Shreeve, T.G., Van Dyck, H., 2003. Towards a functional resource-based concept for habitat: a butterfly biology viewpoint. *Oikos* 102, 417–426.
- Dunning, J.B., Danielson, B.J., Pulliam, H.R., 1992. Ecological processes that affect populations in complex landscapes. *Oikos* 65, 169–175.
- Dynesius, M., Zinko, U., 2006. Species richness correlations among primary producers in boreal forests. *Divers. Distrib.* 12, 703–713.
- Fahrig, L., 2001. How much habitat is enough? *Biol. Conserv.* 100, 65–74.
- Graham, M., 2003. Confronting multicollinearity in ecological multiple regression. *Ecology* 84, 2809–2815.
- Gauchi, J.-P., Chagnon, P., 2001. Comparison of selection methods of explanatory variables in PLS regression with application to manufacturing process data. *Chemometr. Intell. Lab. Syst.* 58, 171–193.
- Holt, R.D., 1997. From metapopulation dynamics to community structure—some consequences of spatial heterogeneity. In: Hanski, I.A., Gilpin, M.E. (Eds.), *Metapopulation Biology—Ecology, Genetics, and Evolution*. Academic Press, San Diego, USA, pp. 149–164.
- Krauss, J., Steffan-Dewenter, I., Tscharntke, T., 2003. How does landscape context contribute to effects of habitat fragmentation on diversity and population density of butterflies? *J. Biogeogr.* 30, 889–900.
- Legendre, P., Legendre, L., 1998. *Numerical Ecology*, second English edition. Elsevier, Amsterdam.
- Maes, D., Van Dyck, H., 2005. Habitat quality and biodiversity indicator performances of a threatened butterfly versus a multispecies group for wet heathlands in Belgium. *Biol. Conserv.* 123, 177–187.
- Manly, B., 1997. *Randomization and Monte Carlo methods in biology*. Chapman & Hall, London.
- Martens, H., 2001. Reliable and relevant modelling of real world data: a personal account of the development of PLS Regression. *Chemometr. Intell. Lab. Syst.* 58, 85–95.
- McIntosh, A., Bookstein, F., Haxby, J., Grady, C., 1996. Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage* 3, 143–157.
- McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: Spatial pattern analysis program for categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst, available at the following web site: <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.
- Mevik, B.H., 2006. The `pls` package. *R News* 6, 12–17.
- Mevik, B.H., Wehrens, R., 2007. The `pls` package: principal component and partial least squares regression in R. *J. Stat. Soft.* 18, 1–24.
- Nguyen, D., Rocke, D., 2002. Tumor classification by Partial Least Squares using microarray gene expression data. *Bioinformatics* 18, 39–50.
- Novotný, V., Basset, Y., 2000. Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos* 89, 564–572.
- Pearman, P.B., Penskar, M.R., Schools, E.H., Enander, H.D., 2006. Identifying potential indicators of conservation value using natural heritage occurrence data. *Ecol. Appl.* 16, 186–201.
- Pimm, S.L., Russell, G.J., Gittleman, J.L., Brooks, T.M., 1995. The future of biodiversity. *Science* 269, 347–350.
- Prendergast, J.R., Eversham, B.C., 1997. Species richness covariance in higher taxa: empirical tests of the biodiversity indicator concept. *Ecography* 20, 210–216.
- R Development Core Team, 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, <http://www.R-project.org>.
- Ricklefs, R.E., 1987. Community diversity: relative roles of local and regional processes. *Science* 235, 167–171.
- Rossi, J.-P., Mathieu, J., Cooper, M., Grimaldi, M., 2006. Soil macrofaunal biodiversity in amazonian pastures: matching sampling with patterns. *Soil Biol. Biochem.* 38, 2178–2187.
- Schindler, S., Poirazidis, K., Wrba, T., 2008. Towards a core set of landscape metrics for biodiversity assessments: a case study from Dadia national park, Greece. *Ecol. Indicators* 8, 502–514.
- Schmidt, I., Zerbe, S., Betzin, J., Weckesser, M., 2006. An approach to the identification of indicators for forest biodiversity—the Solling mountains (NW Germany) as an example. *Restor. Ecol.* 14, 123–136.
- Schneider, C., Fry, G.L.A., 2001. The influence of landscape grain size on butterfly diversity in grasslands. *J. Insect Conserv.* 5, 163–171.
- Shmida, A., Wilson, M.V., 1985. Biological determinants of species diversity. *J. Biogeogr.* 12, 1–20.
- Similä, M., Kouki, J., Mönkkönen, M., Sippola, A.L., Huhta, E., 2006. Co-variation and indicators, of species diversity: can richness of forest-dwelling species be predicted in northern boreal forests? *Ecol. Indicators* 6, 686–700.
- Turner, M., Gardner, R.H., O'Neill, R.V., 2001. *Landscape Ecology in Theory and Practice*. Springer-Verlag, New York.
- Van Halder, I., Barbaro, L., Corcket, E., Jactel, H., 2008. Importance of semi-natural habitats for the conservation of butterfly communities in landscapes dominated by pine plantations. *Biodivers. Conserv.* 17, 1149–1169.
- Weibull, A.-C., Bengtsson, J., Nohlgren, E., 2000. Diversity of butterflies in the agricultural landscape: the role of farming system and landscape heterogeneity. *Ecography* 23, 743–750.
- Whittaker, R.H., 1972. Evolution and measurement of species diversity. *Taxon* 21, 213–251.
- With, K.A., King, A.W., 1999a. Dispersal success on fractal landscapes: a consequence of lacunarity thresholds. *Landscape Ecol.* 14, 73–82.
- With, K.A., King, A.W., 1999b. Extinction thresholds for species in fractal landscapes. *Conserv. Biol.* 13, 314–326.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab. Syst.* 58, 109–130.
- Wolters, V., Bengtsson, J., Zaitseva, A.S., 2006. Relationship among the species richness of different taxa. *Ecology* 87, 1886–1895.
- Wu, J., 2004. Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecol.* 19, 125–138.
- Zaccarelli, N., Riitters, K., Petrosillo, I., Zurlini, G., 2008. Indicating disturbance content and context for preserved areas. *Ecol. Indicators* 8, 841–853.
- Zonneveld, I.S., 1995. Vicinism and mass effect. *J. Veg. Sci.* 6, 441–444.