Methods in ecology and evolution

Application

SILand : an R package for estimating the spatial influence of landscape on ecological data

UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, 78850, Thiverval-Grignon, France

Abstract

1. The spatial distribution of species is both influenced by local variables and by surrounding characteristics of landscape. Estimating spatial influence of landscape is difficult since it requires estimating the intensity and the spatial scale effects of the numerous landscape variables.
2. Here we present the R package 'SILand' for analyzing geolocated specie punctual observations associated to a landscape description (polygons in shapefile format). The package provides estimatestheboth local and landscape variable effects. Scale of landscape variables effects are modelled over the landscape using spatial influence functions. This generic tool deals with different type of observations (continue, discrete, proportion) and with fixed and random local effects.
3. ‘SILand’ can test significance of local and spatial variables, create map of their relative effects and compute AIC and BIC criteria for modelcomparisons.

We use a case study with carpocapse in a landscape composed by organic and conventional orchards to demonstrate thefunctionality of SILandillustrates the main steps of landscape analysis

Introduction.

Several studies suggest that the distribution of species can depend on both local and landscape variables For example, studies have highlighted that…

Butterfly richness and abundance

depend independently on local cultivation practices (e.g.conventional

vs.organicagriculture) and the landscape context (e.g.proportion oButterfly richness and abundancedepend independently on local cultivation practices (e.g.conventional

vs.organicagriculture) and the landscape context (e.g.proportion of organic fields) (Rundlof et al. 2008).f organic fields) (Rundlof et al. 2008).

However studying relationship between landscape and specie distribution remains challenging because the shape and the scale of landscape effects are unknown in advance of the study(Miguet et al. 2016).

To identify the scale of effect, the common approach consists in creating new summary landscape variables. A characteristic of the species distribution is indeed observed at geolocated punctual sites, named ecological response hereafter. A set of potential scales of effect are a priori choosen. New landscape variables are created by computing measures of a landscape variable around observation siteswithin buffers sizes of each potential scales. A regression model is then applied to link the ecological response to the new landscape buffer variables (linear model, forest ,glm …ref).

For each landscape variable, the scale of effects is determined by the buffer size of the new variable explaining best the ecological response.

One major drawback of this method is to increase artificially the number of explanatory variables. Their number is indeed multiplied by the number of potential scales considered.One then has to face a complex statistical dilemma, dealing with numerous explanatory variables which by construction are highly correlated.Consequently, the chosen potential scales are often few (on average four per landscape variable) and explore a too narrow range (Jackson and Fahrig 2015). Finally the results are obtained considering that the effect of a landscape variable is uniform within the buffer and becomes abruptly zero outside it (Chandler and Hepinstall-Cymerman 2016), even though modelling landscape influence using threshold step function is unrealistic and unjustified biologically (Moilanen and Hanski 2001).

Citer Remm et al. Pour exemple d’application sur les écureuils volants

Several new methods based on distance weighting effects have been proposed to model a distance-decreasing effect and allowed to quantify the scale of landscape effects without an a priori choose of potential scales (Chandler and Hepinstall-Cymerman 2016), (Aue et al. 2011)(Walsh and Webb 2014))(Serckx et al. 2016). However none of these methods are directly available in a ready-to-use software. Here we present “SILand” a package for R statistical computing environment designed for provide a generic tool for the estimation of landscape effects on an ecological response. Using the framework proposed by Chandler et al. 2016, the main functions of SILand allow the user (i) to estimate the intensity of local and landscape effects and the parameters of the spatial influence function (SIF) describing the effect scale, (ii) to test these effects and (iii) to compare models. Several types of ecological response (continue, discrete, proportion) are handled. Local variables can be traited as fixed or random effects. In addition, the package provided visualization tool to plot the estimated spatial and local effectson the observations map.t. We exemplified the package use byanalyzing the effect of conventional and organic apple tree orchards on the density of carpocapse.

.ref à lire (Mayor, Schneider, Schaefer, & Mahoney,

2009; Urban, 2004; Wheatley, 2010). The

Urban, D. L. (2004). Modeling ecological processes across scales. Ecology,

86, 1996–2006.

Wheatley, M. (2010). Domains of scale in forest-landscape metrics:

implications for species-habitat modeling. Acta Oecologica, 36,

259–267.

Mayor, S. J., Schneider, D. C., Schaefer, J. A., & Mahoney, S. P. (2009).

Habitat selection at multiple scales. Ecoscience, 16, 238–247.

Henry et al., 2012Henry, M., Fröchen, M., Maillet-Mezeray, J., Breyne, E., Allier, F., & Odoux,

J.-F. (2012). Spatial autocorrelation in honeybee foraging activity

reveals optimal focus scale for predicting agro-environmental scheme

efficiency. Ecological Modelling, 225, 103–114.

Model

The estimation of local and landscape effects is based on framework proposed by Chandler et al. (2016). The estimation of these effects require a discretization or rasterization of the landscape, that is the landscape is represented by a regular grid. Each pixel (or cell) of this grid is characterized by a set of landscape covariates and an ecological response measured at different sites. The ecological response is denoted by for .., denotes the value (continuous or discrete) for local variable measured at site and for variable , and with the set of local variables. For pixel , denotes the value (continuous or discrete) for landscape variable with the set of landscape variables. Classicaly, =1 indicates that the landscape variable is observed in pixel , and 0 otherwise.

For simplicity, we suppose that the ecological response is modeled with a Gaussian distribution with where is the intercept, and are parameters associated to local and landscape variables for and . represents the set of pixels for the grid. The unknown scale at which the landscape variable affects the ecological response is considered through the spatial influence function This function represents the effect of pixel on observation at site , and this effect is decreasing with Euclidean distance between the location of the site and the location of pixel . The scale of landscape variable is defined through the parameter. Parameters are estimated by maximizing the likelihood and the maximizing likelihood is based on an iterative procedure using the fact that conditionally to the scale parameter the model is linear.

By maximizing likelihood, tests on local and landscape variables, comparison models with AIC or BIC criteria or different shapes (Gaussian or Exponential) of the spatial influence function can be performed. The package can handle a large variety of models. For example, Poisson and Binomial distribution can be considered for counting or presence-absence data. In the part concerning the local effects, mixed models can be considered to take into account repeated measurements.

CASE STUDY

We demonstrate the capabilities of SILand on an example with codling moths, an insect pest specialized on orchards, previously described and analyzed in Ricci et al. 2009. The codling moths data consists in repeated measurements of number of the overwinter larvae observed in 76 geolocalised orchards over a 70-km² area in southeastern France. The landscape data contains the locations of orchards of the area associated to their characteristics pear vs. apple, organic vs. conventional orchards (figure associée au jeux de données).

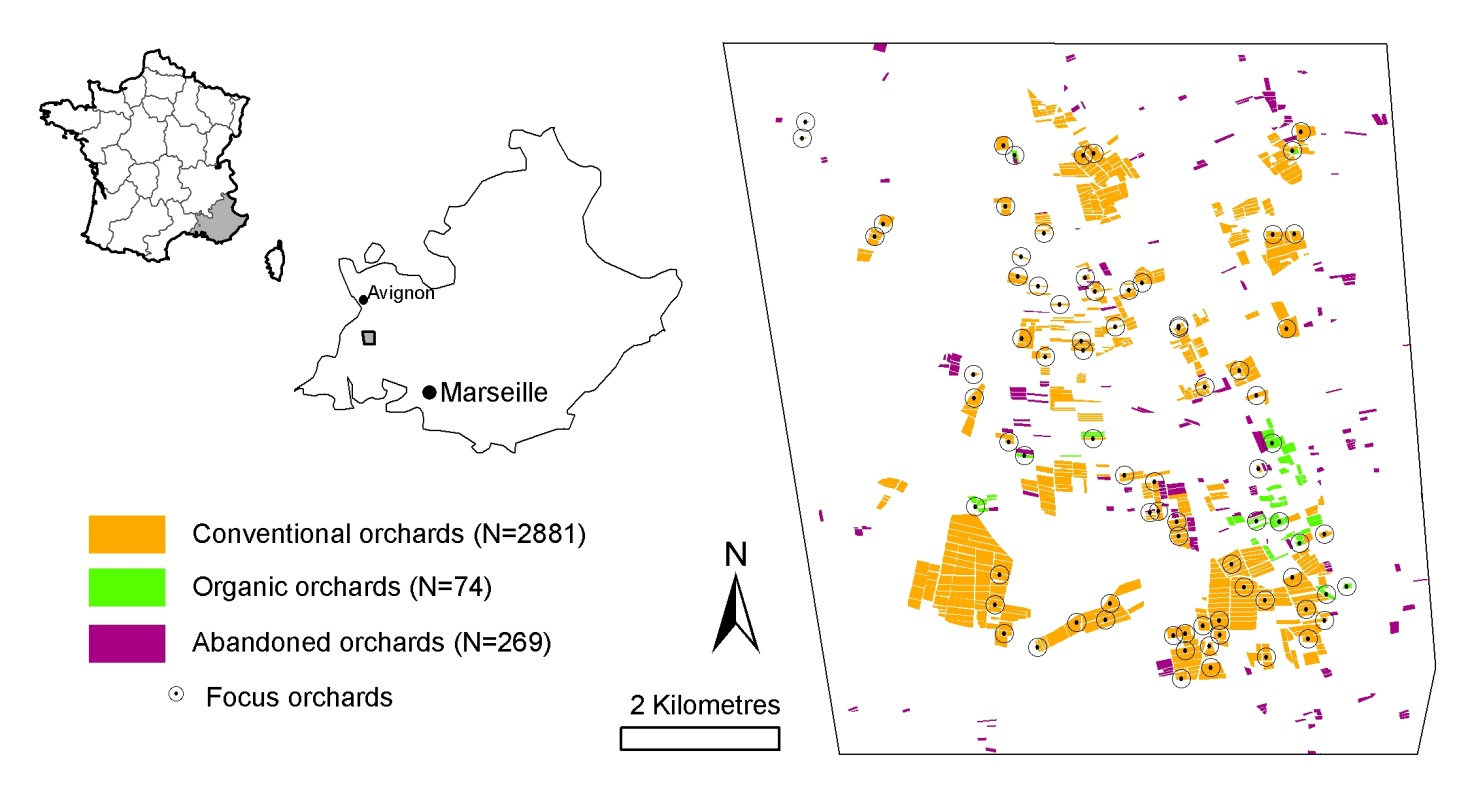


Figure 1

1.Data loading from sig files

We begin by loading data concerning the landscape and the ecological reponse. This step is important since the format of data input for siland function has to have a precise structure. A straightforward way is to work with sig files and to use the functions data.gis.siland() and land.gis.siland(). The structure of objects R for the landscape description and the ecological response are respectively detailed in help(landSiland) and help(dataSiland).

dataC=data.gis.siland(path,layerdata,yname,locvarname,as.factor.locvar=NULL)

landC=land.gis.siland(path,layerland,landvarname,vallandvar,wd=100)

2. Fitting model with local and landscape variables

Model estimation is fitted with the function island that needs different options.

resG=siland(loc.model= carp~BioConv,land=landC,data=dataC,sif="exponential",initSIF=c(100,100),test=T,family="poisson")

The argument loc.modelspecifies the part of model concerning local varibles. The syntax of this argument is similar to the one used for function lm() or glm(). The influence of landscape is modeled by the elemets of the list landC. In this example, the list landC have two components named Bio and Conv. So, in this model the objective is to estimate the influence of surrounding organic and conventional features on the ecological response. The sif argument corresponds to the shape of the influence function (Gaussian or Exponentiel), initSIF specifies the initialization for the two mean distance of the sif function.

The coefficients obtained from summary(resG) are the estimated parameters for local and landscape variables. The values for SIF.Bio and SIF.conv correspond respectively to the parameter of the spatial influence function for landscape variables Bio and Conv. Hereafter, we present in details the parametrization of this function.

The argument test indicates if tests for parameters have to be computed or not. Since tests are based on likelihood ratio test with embedded models, tests can be time consuming if a large number of variables are considered. Finally, the argument family gives the assumed distribution of ecological response.

A comparison between the full model (model with local and landscape variables) and a local model (only local variables) is realized. The AIC for the full model, the local model and pvalue from likelihood ratio test are computed.

>summary(resG)

Coefficients:

(Intercept) BioConv Bio Conv SIF.BioSIF.Conv

3.424345 -1.005986 11.217203 -12.104822 416.021464 803.565290

pvalue (L.R. Test):

BioConv Bio Conv

<1e-16 <1e-16 <1e-16

AIC: 9611.82 AIC (no landscape): 13771.59

(No landscape effect) p-value: <1e-16

The argument loc.modelcan include random effects by using the syntax from package lme4. If several measurements have been realisedine the orcjard of the study, one can estimate a random orchard effect

>resRandom=siland(loc.model= carp~BioConv+(1|orchard),land=landC,data=dataC,sif="exponential",initSIF=c(100,100),test=T,family="poisson")

>summary(resRandom)

3. Selection of the shape of influence for landscape variables

The exponential and Gaussian shape can be compared with the AIC criterion. The exponential and gaussian shape are both defined by one parameter denoted. In the Exponential case, and for the Gaussian case, where represents the Euclidean distance between an observed siteand a pixel. One can see the function like a density function that gives decreasing weights to pixels with the distance from the observed site. Using these parametrization, parameter can be interpreted as an expected weighted distance for the influence of pixel.If is high, a pixel of the landscape has an influence at a greater distance than if is low.

resG=siland(loc.model=CarpoVivan~1,land=D$dataland,data=D$datatable,initSIF=c(20,20),sif=”gaussian”,family=”poisson”)

summary(resG)

resE=siland(loc.model=CarpoVivan~1,land=D$dataland,data=D$datatable,initSIF=c(20,20),sif=”exponential”,family=”poisson”)

summary(resE)

4. Graphic representation

READING DATA

Shapefile format

Type of object

ESTIMATION AND TESTS OF EFFECTS

Siland

Options

Results maximum de vraisemblance

Estimation

Tests

Testing effects

Nested models spatial effects

COMPARING MODELS

AIC

BIC

Functions of influence

CREATING MAPS OF LANDSCAPE EFFECTS

Aue, Birgit, Ekschmitt Klemens, Hotes Stefan, and Wolters Volkmar. 2011. “Distance Weighting Avoids Erroneous Scale Effects in Species‐habitat Models.” *Methods in Ecology and Evolution* 3 (1): 102–11. https://doi.org/10.1111/j.2041-210X.2011.00130.x.

Chandler, Richard, and Jeffrey Hepinstall-Cymerman. 2016. “Estimating the Spatial Scales of Landscape Effects on Abundance.” *Landscape Ecology* 31 (6): 1383–94. https://doi.org/10.1007/s10980-016-0380-z.

Jackson, Heather Bird, and Lenore Fahrig. 2015. “Are Ecologists Conducting Research at the Optimal Scale?” *Global Ecology and Biogeography* 24 (1): 52–63. https://doi.org/10.1111/geb.12233.

Miguet, Paul, Heather B. Jackson, Nathan D. Jackson, Amanda E. Martin, and Lenore Fahrig. 2016. “What Determines the Spatial Extent of Landscape Effects on Species?” *Landscape Ecology* 31 (6): 1177–94. https://doi.org/10.1007/s10980-015-0314-1.

Moilanen, Atte, and Ilkka Hanski. 2001. “On the Use of Connectivity Measures in Spatial Ecology.” *Oikos* 95 (1): 147–51. https://doi.org/10.1034/j.1600-0706.2001.950116.x.

Serckx, Adeline, Marie-Claude Huynen, Roseline C. Beudels-Jamar, Marie Vimond, Jan Bogaert, and Hjalmar S. Kühl. 2016. “Bonobo Nest Site Selection and the Importance of Predictor Scales in Primate Ecology.” *American Journal of Primatology* 78 (12): 1326–43. https://doi.org/10.1002/ajp.22585.

Walsh, Christopher J., and J. Angus Webb. 2014. “Spatial Weighting of Land Use and Temporal Weighting of Antecedent Discharge Improves Prediction of Stream Condition.” *Landscape Ecology* 29 (7): 1171–85. https://doi.org/10.1007/s10980-014-0050-y.

Model

Y ~ L(local+spatial)

Modélisation of spatial effect :

FIS : concept, parameterinterpretation  
 family

Equation des contributions spatiales

Modélisation of local effect : fixed or random effect

Type of data : L different family (mv abund)

Importing Data

* data<-sig

return objet de type de machin de forme

polygone = point width