



# **Detecting threatened biodiversity by urbanization at regional and local scales using an urban sprawl simulation approach: Application on the French Mediterranean region**

Ruppert Vimal, Ghislain Geniaux, Pascal Pluvinet, Claude Napoléone, Jacques Lepart

## **► To cite this version:**

Ruppert Vimal, Ghislain Geniaux, Pascal Pluvinet, Claude Napoléone, Jacques Lepart. Detecting threatened biodiversity by urbanization at regional and local scales using an urban sprawl simulation approach: Application on the French Mediterranean region. *Landscape and Urban Planning*, 2012, 104 (3-4), pp.343 - 355. <10.1016/j.landurbplan.2011.11.003>. <hal-02456117>

**HAL Id: hal-02456117**

**<https://univ-tlse2.hal.science/hal-02456117v1>**

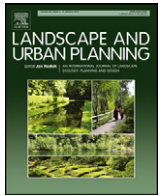
Submitted on 27 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



# Detecting threatened biodiversity by urbanization at regional and local scales using an urban sprawl simulation approach: Application on the French Mediterranean region

Ruppert Vimal<sup>a</sup>, Ghislain Geniaux<sup>b,\*</sup>, Pascal Pluvinet<sup>a</sup>, Claude Napoleone<sup>b</sup>, Jacques Lepart<sup>a</sup>

<sup>a</sup> Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175, CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France

<sup>b</sup> Ecodéveloppement UR 767, INRA, Domaine St Paul, 84914 Avignon Cedex 9, France

## ARTICLE INFO

### Article history:

Received 10 February 2011

Received in revised form 18 October 2011

Accepted 11 November 2011

Available online 15 December 2011

### Keywords:

Biodiversity

Prospective

Urbanization

Land-use change model

Mediterranean region

## ABSTRACT

Urbanization is a major threat on biodiversity due to the direct destruction of natural and semi-natural habitats and to the indirect impacts caused by urban areas beyond their limits. In this study, we proposed a methodological framework to assess the potential impacts of current and future urbanization on high diversity sites and their associated species across the entire French Mediterranean region. Using a 100 m grid cell, we adapted a land-use change (LUC) model to project future urbanization over a 20 years period. We developed a multi-level approach based on three impacts of urban development: the direct consumption of high diversity sites, the indirect urban effects on the surrounding area over a scale of 2 km and a scale of 50 km. Our model predicts that 35% of the total number of sites can potentially be impacted by urbanization (i.e. at least 1 ha predicted to be built between 2006 and 2030). 3.2% of the total number of sites may lose 10% or more of their surface area to urbanization. We found that the impacts of urban area and urban growth differ among the three different pressure indicators in terms of surface area and localization of sites, number and nature of species impacted and variation of these patterns between the two dates. In general, most of the sites under pressure are located in the coastal part of the study region and are of smaller surface area than average.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Rapid human population growth over the last century has resulted in urban areas covering about 2% of total land surface (Grimm, Grove, Pickett, & Redman, 2000). As human populations continue to grow, major urban areas will continue to expand (Meyer & Turner, 1992). Although the concentration of people in major cities and the densification of new constructions can help to protect natural and agricultural resources elsewhere (Forman, 2008), urbanization represents a major threat to biodiversity across the world (Chapin et al., 2000; McDonald, Kareiva, & Forman, 2008; Wilcove, Rothstein, Dubow, Phillips, & Losos, 1998).

It is of major concern for biodiversity conservation for several reasons. First, urbanization is among the least reversible of land-use change. Second, because urbanization is often concentrated in areas of high net primary production which are also areas with very high

species richness (Araujo, 2003; Balmford et al., 2001; Luck, 2007a; Vazquez & Gaston, 2006), its effects on biodiversity are much greater than if it were randomly distributed in space. Third, urban and suburban areas facilitate the presence of exotic species which thrive in habitats where human activities have removed the native dominant species. Hence, the number of non-native species is high and the contribution of urban (and suburban) areas to the conservation of global biodiversity is very low (McKinney, 2002). Finally, urban impacts on biodiversity can extend far beyond the city limits (Forman, 2008; Luck, Jenerette, Wu, & Grimm, 2001). Indeed, urban areas threaten ecosystems as a result of both direct habitat conversion (Clergeau, Savard, Mennechez, & Falardeau, 1998; McKinney, 2002) and through various indirect effects, e.g. land use change in the periphery of urban area, fragmentation of the territory by linear infrastructures associated with communication and transport among urban areas, waste generation and water pollution, and disturbances associated with recreational activities around urban areas (DeFries, Hansen, Turner, Reid, & Liu, 2007).

There is thus an urgent need to pay close attention to the spatial distribution of urban areas and to project its evolutions in the future in relation to the distribution of biodiversity conservation interests. The potential impacts of urban spatial expansion on biodiversity have been studied across a diversity of scales ranging

\* Corresponding author. Tel.: +33 0 4 32 72 25 64; fax: +33 0 4 32 72 25 62.

E-mail addresses: [geniaux@avignon.inra.fr](mailto:geniaux@avignon.inra.fr) (G. Geniaux), [ppluvinet@yahoo.fr](mailto:ppluvinet@yahoo.fr) (P. Pluvinet), [cnapoleone@avignon.inra.fr](mailto:cnapoleone@avignon.inra.fr) (C. Napoleone), [jacques.lepart@cefe.cnrs.fr](mailto:jacques.lepart@cefe.cnrs.fr) (J. Lepart).

from the international and national level scale (e.g. Jenerette & Potere, 2010; Theobald & Romme, 2007) to regional level effects, mostly associated with habitat fragmentation (Lawson, Lamar, & Schwartz, 2008; Manley, Parks, Campbell, & Schlesinger, 2009; Martinuzzi, Gould, & Gonzalez, 2007), and impacts observed in and around particular urban centers (Jarrige, 2004; Wu, Hu, He, Bu, & Xi, 2007). One of the recurrent difficulties is to assess the relative effects of the diversity of impacts caused by urbanization because it can range from destruction of habitats to indirect effects associated with pollution or noise caused by human proximity. Global analyses have often failed to support the multi-scale dimensions of urban threats on biodiversity and suffer from a lack of precise data on both socio-economic conditions and biodiversity inventories.

Over the past decades, several models have been developed to project and quantify future land use and land cover for ecosystem impact assessment (Irwin & Geoghegan, 2001; Parker, Manson, Janssen, Hoffmann, & Deadman, 2003; Veldkamp & Lambin, 2001; Verburg, Schot, Dijst, & Veldkamp, 2004), mainly in land-use-change (LUC) models. LUC models aim to show how or where irreversible changes could arise in the future, in order to adapt current public policy (Conway & Lathrop, 2005; Lambin, 1997). Urbanization can be modeled through various approaches (Irwin, 2010) from complex descriptions of urban expansion with many parameters describing multiple levels of land use change (Alberti, 1999; Fontaine & Rounsevell, 2009; Landis, 1995), to simplified models using a minimal set of parameters on a large scale (Batty, 1991; Fagan, Meir, Carroll, & Wu, 2001; Rouget, Richardson, Cowling, Lloyd, & Lombard, 2003). In general, most of the fine-scale modeling studies have only been implemented for a single city (e.g. Jenerette & Potere, 2010; Taylor, Brown, & Larsen, 2007). A similar modeling framework at regional scale would allow a more precise understanding of the interactions between micro-level parameters and macro-level land use change and thus help assess the potential impacts of future urban spread on biodiversity.

The aim of this manuscript is both to provide (i) a new urban sprawl model at fine scale for the entire Mediterranean region of southern France and (ii) a methodological approach to assess urbanization impacts on biodiversity: what kind and where would be the most endangered species if urban trend changes were likely to continue as in the last 20 years?

## 2. Methods

### 2.1. The overall methodological approach

The methodological approach of the study can be partitioned in two parts: First, future urbanization was forecasted at 20 years horizon with 100 m grid cell. We used vectorized data on housings and land use plan combined with a statistical hierarchical model at two levels that projects number of housings at communal level and localizes the housings leading to land conversion at cell level.

Second, we assessed the potential impacts of current and future urbanization on high biodiversity sites and their associated species. In order to account for a range of different urban threats, we developed a multi scale approach based on three urban pressure indicators from precise urban arrangements on a local scale, to global human density impacts over a larger scale. To do so, we examined urban development and its potential impact in three ways namely the direct consumption of high diversity sites, and indirect urban effects on the surrounding area over a scale of either 2 km or 50 km.

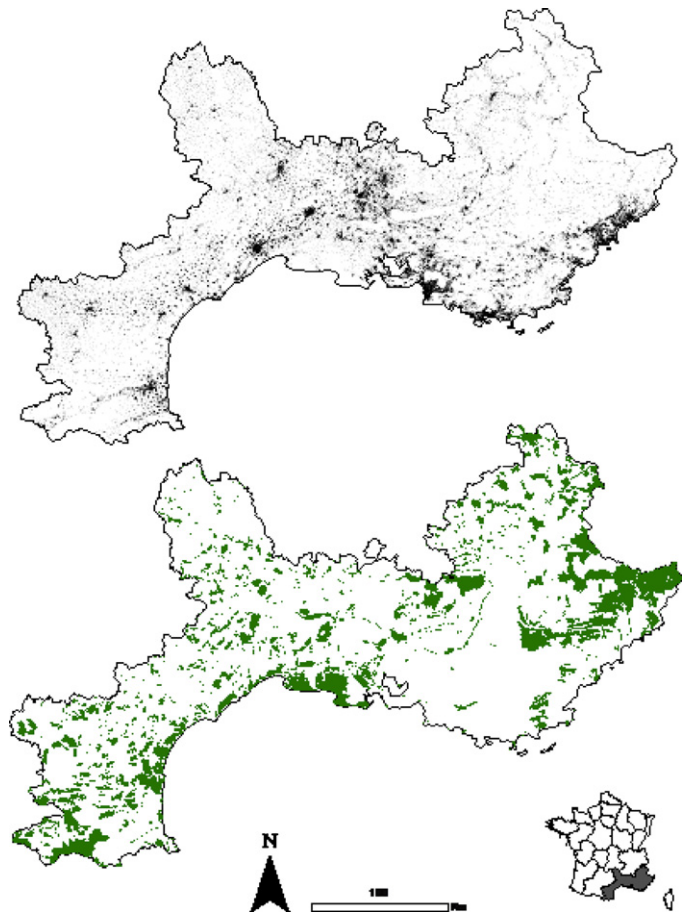


Fig. 1. The Mediterranean region of southern France with (a) the current distribution of urbanized areas and (b) the sites of high biodiversity interest (ZNIEFF inventory).

### 2.2. Study area

The study region covers 59 660 km<sup>2</sup> in the Mediterranean region of southern France (Fig. 1). It represents two regional administrations: the *Languedoc-Roussillon* (LR) region between the western part of the Rhône valley and the Spanish border and the *Provence-Alpes-Côte d'Azur* (PACA) region between the eastern and the Italian border. These two regions have 11 administrative subdivisions (counties or “*départements*”) and 2508 municipalities, each of which has its own local land use plan. The study region is one of the major World hotspots of biodiversity (Medail & Quezel, 1999; Myers, Mittermeier, Mittermeier, da Fonseca, & Kent, 2000; Shi, Singh, Kant, Zhu, & Waller, 2005), but is also one of the most transformed regions with a marked landscape diversity due to a long history of human land-use – notably extensive agro-sylvo-pastoral practices and cultivation (Blondel, Aronson, Bodiou, & Boeuf, 2010; Thompson, 2005). The main landscape types which occur in this region are coastal landscapes with lagoons, marshes, cliffs and dunes, lowland “*garigues*” often as a mosaic landscape with cultivated areas, vast areas of vineyards, extensive upland plateau areas, and hilly or mountainous landscapes in the southern tip of the Massif Central, the south-eastern Pyrenees and the pre-Alps. The Mediterranean region is currently undergoing massive coastal urbanization and infrastructure development (Benoit & Comeau, 2005). The French Mediterranean region has currently the highest population growth of France (Bessy-Pietri, 2000), particularly in lowland areas close to the coast. This led to rapid spread of urbanization around towns and villages as well

as infrastructure development along the coasts associated with tourism.

### 2.3. Data

#### 2.3.1. Biodiversity data

As part of the national inventory of high ecological value sites (“Zones Naturelles d’Intérêt Ecologique Faunistique et Floristique” or ZNIEFF), a list of determinant species of conservation interest has been elaborated in each region of France and high diversity sites based on the presence of these species designated. This inventory is validated by a regional scientific council (“Conseil Scientifique Régional du Patrimoine Naturel” or CSRPN). We used the inventory of the two administrative regions as a basis for our study. To identify the list of determinant species, regional specialist organizations weighted and noted each species of a given taxonomic group to define their conservation interest (INPN, 2006). The main criteria used for this purpose were local rarity, quantified by the number of distinct localities where a species has been recorded in the region, and the regional responsibility, which is estimated by the ratio between the species frequency in the region and its frequency in France. Additional criteria such as international, national, or regional protection status were also considered. For the purpose of this study we used taxonomic groups for which we had sufficient information on the regional distribution and abundance; 1040 vascular plants, 28 mammals, 20 herptiles (reptiles and amphibians), 86 birds and 16 fishes. Following the inventory of these species, 1315 high diversity sites (18.2% of the surface of the study region) were designated and delimited by the regional operators (Fig. 1). Only species reproducing on a given site were considered to be present and minor adjustments made according to species composition based on expert consultation.

#### 2.3.2. Urbanization data

For the whole study area, we used four kinds of databases:

- The 2008 “built-up” layer of BD TOPO®/RGE Geodatabase (IGN Institut Géographique National) which is the topographic component of French RGE (literally “Frame of reference at large scale”). This layer contains around 300 000 polygons for each administrative subdivision. We used the aggregated BD TOPO® data for undifferentiated building. We considered a grid cell as built-up if at least 2.5% of its surface was covered by built-up polygons.
- The Land Use Plan Geodatabase was obtained from the two nationally administered regional environmental agencies (*Direction Régionale de l’Environnement, de l’Aménagement et du Logement*), which are in charge of the legal control of town and country planning policy. They have a digitalized land-use plan (LUP) for each municipality. These LUPs are harmonized for comparative use. The resulting regional geodatabase is called a “Generalized LUP” and is updated every two years. For the purpose of this study, we reclassified LUPs into three types of zoning, indexed hereafter by *z*: high urban density zoning (URB), future Urban/Activities zoning (FURB), isolated houses with agricultural/natural zoning (NONBDEV). Since LUPs do not exist for most rural and uninhabited municipalities, we used the Corine Land Cover Database (CLC 2006) to define three zones similar to those defined above. Dense urban zones of CLC 2006 have been classified as URB zonings, other urban classes of CLC as FURB zoning and the rest as NONBDEV zoning. While this method appears inelegant, it remains coherent with the philosophy of urban regulation law (called MARNU) in non zoned municipalities. Moreover, it represents only 3% of the regional housings and 27% of the total number of municipalities.
- In order to tune some parameters of the simulation model (see Section 2.4.3 on “Hierarchical links between models”), we used

a cadastral vector database available for the PACA region only. This database contains two important pieces of information. First, it denotes parcel boundaries (3 millions polygons) and house delimitation (1.6 millions polygons; “Plan Cadastral Informatisé”). Second, it provides house and owner characteristics (fiscal database commonly called MAJIC II) that allows to identify for each residential house the date of construction (Geniaux, Podjleski, & Leroux, 2009).

- At municipal level, we used census data for population, housing, activities and employment (INSEE census data). Following Geniaux, Dumas, Bartoli, Cezanne-Bert, and Napoléone (2005) and Geniaux (2010), we built a classification of municipalities to describe statistically urban density dynamics and LUP policies. This classification accounted for a classification of total houses per municipality and a classification of the population in the 50 km neighborhoods of the municipal surroundings. Total housings per municipality were classified into 4 levels (0–200, 200–2000, 2000–10 000, >10 000). Surroundings population for each municipality was estimated using a Gaussian kernel weighting based on the distance in km between municipalities – with a 10 km bandwidth – and classified into 4 levels using quartiles. It resulted in 11 effective levels indexed hereafter by *c*.
- Information from these four databases has been reported on a grid with a cell size of 100 m representing 6 millions of cells for the two regions.

### 2.4. Urban model

Our urban model aims at forecasting land use change (LUC) up to a 20 years time-horizon using the 100 m grid cell. The conceptual framework of the urban model is illustrated in Fig. 2. It can be classified as a hierarchical one because it combines two statistical models at different levels. At communal scale, a first level model was used to forecast municipal number of housings in the next 20 years based on demographic and urban data. At the scale of a cell of 100 m × 100 m, a second level model was used to localize the new housings in each zone of the land use plan. To do so, we determined likelihood for cells not yet built to be built based on a logit estimate. Three local urban parameters were necessary to estimate the potential range of housings changes for each land use plan zone *z* in each municipality class *c* in order to link hierarchically both models. These parameters were obtained from a statistical analysis of building dynamics at parcel level for the PACA region for the period 1990–2005.

Our modeling framework distinguishes itself from previous works on LUC model by using information on land use plan for more than 1500 municipalities. As noted by Irwin (2010), almost all LUC models do not use geographic data on land use plan and are based on a theoretical framework in which land market and land use are not regulated by public institutions.

The logit probability is estimated in each zoning type (URB, FURB and NONBDEV) because in 20 years a large share of the new houses will be localized in new urban zones (currently zoned as non developable) through land use plan changes. Modeling the likelihood of future land use change in non developable zones follows the assumption that “zoning follows the market” (Pogodzinski & Sass, 1994; Thorson, 1994; Wallace, 1988). This assumption supports the idea that high land prices or high probabilities to be built will promote land use plan changes from non developable to developable zoning in the future Geniaux, Ay, and Napoléone (2011).

Our urban sprawl model can be considered as only simulate a trend scenario because parts of its development are based on recent urban dynamics (e.g. the estimated projection of new housings at municipal scale and the assessment of the proportion of new housings per zone *z*). However the likelihood to be built for a non-built cell is calculated without differentiating recent



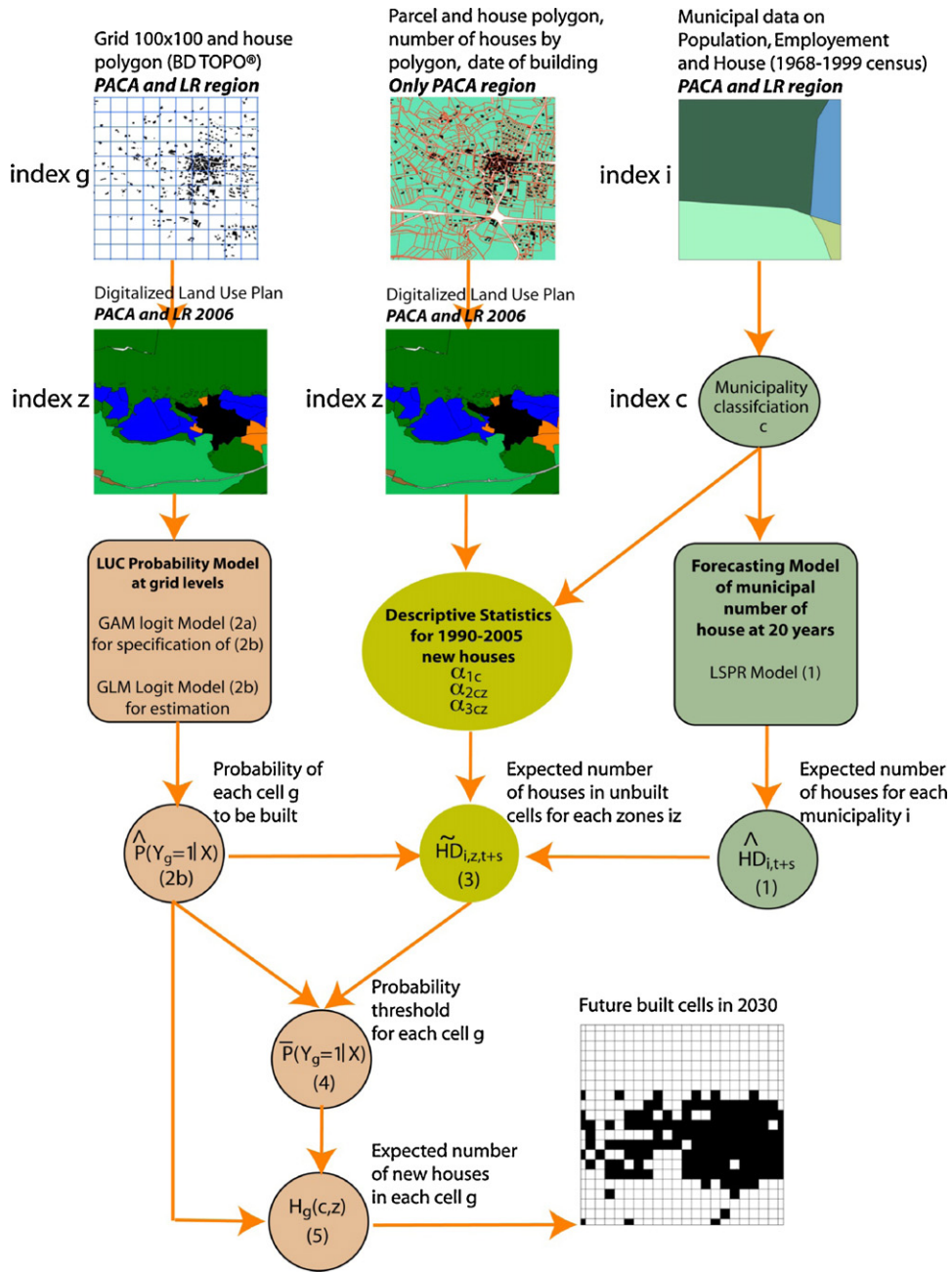


Fig. 2. Overview of the urban simulation model.

housing constructions and less recent housing constructions due to the lack of data on date of construction for urban areas in the region Languedoc-Roussillon.

#### 2.4.1. Projection model for total number of housings in 2030 at municipal scale

At municipal scale, we estimated a projection of the total number of housings for 2030 for each municipality  $i$ , noted  $HD_i$ . Total number of municipal housings can be mainly explained by sequences of past number of housings (we used the four preceding census), by the county, and by inhabitants and employments in the surrounding 50 km radius. Moreover, we used strata in order to have different coefficients values for 7 types of municipalities according to their total number of housings (7 levels of number of

housing of the preceding census noted  $CHD_{i,t-1,k}$ ):

$$HD_{t,i} = \beta_0 + \beta_1 \text{County} + \sum_{k=1}^7 \sum_{s=1}^4 \alpha_{k,s} CHD_{i,t-1,k} \times [HD_{i,t-s} + spop_{t,i} + slab_{t,i}] + \epsilon_i \quad (1)$$

Model (1) was estimated with a weighted least square in which the weights are equal to the inverse of the endogenous lag  $1/HD_{t-1,i}$ . This type of regression is called Least Squares Percentage Regression (LSPR) and aims to minimize the square of the relative errors (Tofallis, 2009). This type of regression is necessary to avoid that large municipalities may have too much influence on the results and lead to high relative errors for small municipalities. Coefficient estimates of (1) allow projecting  $HD_i$  at  $t+s$ .

**Table 1a**  
Covariates of model (1).

Discrete covariates		Frequency
<i>CHD</i> <sub>1999</sub>		
Classes of municipal number of house (main home) in 1999 ( <i>HD</i> <sub>1999</sub> )		
[0, 200] houses		0.388
(200, 500] houses		0.250
(500, 1000] houses		0.143
(1000, 2000] house		0.107
(2000, 5000] houses		0.068
(5000, 15 000] houses		0.029
(15 000, 4 000 000] houses		0.015
<i>ZonePop</i>		
Classes of smoothed surrounding population (50 km, <i>sspop50</i> ) <sup>a</sup>		
Urban core area ( <i>sspop50</i> > 10 000)		0.045
First commuter belt 1 ( <i>sspop50</i> ∈ [5000–10 000])		0.074
Second commuter belt ( <i>sspop50</i> ∈ [1200–5000])		0.254
Rural area ( <i>sspop50</i> < 1200)		0.627
Continuous covariates		Mean and [range]
<i>HD</i> <sub>2006</sub>	Municipal number of house (main home) in 2006	1724 [5–405 600]
<i>HD</i> <sub>1999</sub>	Municipal number of house (main home) in 1999	1567 [5–385 600]
<i>HD</i> <sub>1990</sub>	Municipal number of house (main home) in 1990	1392 [3–369 100]
<i>HD</i> <sub>1975</sub>	Municipal number of house (main home) in 1975	991 [3–352 500]

$$^a \text{sspop50}_i = \sum_j e^{\text{dist}_{ij}^2/10^2} \text{pop}_j.$$

#### 2.4.2. Projection model for probability of land use conversion at cell level

Our model estimates the probability for each cell *g* with no house ( $Y_g = 0$ ) to be developed for residential use ( $Y_g = 1$ ) knowing a set of covariable *X* and the type of LUP zone *z* of each cell *g*. In order to estimate such probability in a spatial and large database context, we used a semi-parametric spatial generalized additive logit model (Geniaux & Napoleone, 2008; Hastie & Tibshirani, 1993; Wood, 2006) to specify a parametric GLM probit.

A Generalized Additive Model (Hastie & Tibshirani, 1990) is an extension of the Generalized Linear Model (GLM) in which the linear predictor is specified as the sum of smooth functions of determining factors. We obtained a geoadditive logit model by integrating, in a smoothing function, longitude *y* and latitude *x*, and by using a logistic distribution:

$$P(Y_g = 1|X_1, X_2, x, y, z) = \beta X_1 g + \sum_m s_m(X_2 g) + s(x_g, y_g) + \epsilon_g \quad (2a)$$

The GAM logit model (2a) has been used (i) to identify non linear-ity in continuous covariables, in particular distance and residential density covariates and (ii) to obtain a final specified model that flattens as possible the spatial smoothed terms. This specification process used various subsamples of 100 000 cells randomly chosen between the 6 millions of cells of the study area.

Non linear covariates  $X_{2g}$  has been split into two or three linear parts and integrated in a full parametric GLM logit that can be expressed as:

$$P(Y_g = 1|X, z) = \beta_0 + \beta_1 \text{clc90} + \beta_2 \text{sl} + \beta_3 \text{zonepop} + \beta_4 \text{nsd} + \beta_5 \text{lpa} \\ + \beta_6 \text{duc1} + \beta_7 \text{duc2} + \beta_8 \text{dsr1} + \beta_9 \text{dsr2} + \beta_{10} \text{pr1} \\ + \beta_{11} \text{pr2} + \beta_{12} \text{dmr1} + \beta_{13} \text{dmr2} + \beta_{14} \text{area} \\ + \beta_{15} \text{apr} + \beta_{16} \text{WYp} + \epsilon_g \quad (2b)$$

Tables 1a and 1b describe labels and descriptive statistics for each covariate of the final model. Due to the very large number of grid cells, model (2b) has been estimated separately in each of the

**Table 1b**  
Covariates of model (2a).

Discrete covariates		Frequency
<i>clc90</i>		
Corino Land Cover 1990 (15 classes)		
Artificial surfaces		0.0406
Agricultural areas		0.2973
Forests and semi-natural areas		0.6280
Wetlands		0.0153
Water bodies		0.0188
<i>sl</i>		
Mean slope of cell (3 levels)		
Under 10%		0.6631
10–30%		0.3071
More than 30%		0.0297
<i>ZonePop</i>		
Classes of smoothed surrounding population (50 km, <i>sspop50</i> ) <sup>a</sup>		
Urban core area ( <i>sspop50</i> > 10 000)		0.0773
First commuter belt 1 ( <i>sspop50</i> ∈ [5000–10 000])		0.0934
Second commuter belt ( <i>sspop50</i> ∈ [1200–5000])		0.2406
Rural area ( <i>sspop50</i> < 1200)		0.5886
<i>nsd</i>		
North/south/flat orientation dominance		
		$N = 0.3406/S = 0.3050/F = 0.36$
<i>lpa</i>		
Intersects a protected area (2 levels)		
		$Y = 0.0509/N = 0.9490$
Continuous covariates		Mean and [range]
<i>duc1</i>	Shortest distance to urban core of a town (<1000)	446.9 [0–1000]
<i>duc2</i>	Shortest distance to urban core of a town ( $\geq 1000$ )	2748 [1000–7348]
<i>dsr1</i>	Shortest distance to all road (<1000)	315.30 [0–1000]
<i>dsr2</i>	Shortest distance to all road ( $\geq 1000$ )	1851 [1000–11 390]
<i>dmr1</i>	Shortest distance to main road (<12 500)	3976 [0–12 500]
<i>dmr2</i>	Shortest distance to main road ( $\geq 12 500$ )	17 370 [12 500–35 300]
<i>pr1</i>	Ratio of road surface (<0.1)	0.006085 [0–0.1]
<i>pr2</i>	Ratio of road surface ( $\geq 0.1$ )	0.1581 [0.1–1]
<i>area</i>	Area of municipality	6359 [0–222 800]
<i>apr</i>	Area/perimeter ratio of municipality	0.8490 [0.2462–9.5790]
<i>WYp</i>	Ratio of built surface in contiguous cells	0.01093 [0–1]

$$^a \text{sspop50}_i = \sum_j e^{\text{dist}_{ij}^2/10^2} \text{pop}_j.$$

11 administrative sub-regions (French county or “department”) of the studied area (around 500 000 cells in each county). Moreover, model (2b) has also been estimated separately for each type of zoning *z* and for each county in order to account for the heterogeneity of urbanization process and density between types of zones *z*. This led to 33 logit models estimation. Finally, we obtained for each cell *g* the probability to be developed before 2030 noted  $\hat{P}(Y_g = 1|X, z)$ .

#### 2.4.3. Hierarchical links between models (1) and (2b)

In order to identify on an infra-municipal level the precise threshold for realization of land use change probability for each type of land use plan zones *z* noted  $\hat{P}(Y_g = 1|X)$ , we used statistics on housing at parcel scale by type of zoning *z* and by type of municipality *c* for recently built houses (after 1990) in PACA region for which the date of construction is known.

- $\alpha_{1c}$  is the proportion, for each type of municipality *c*, of new housings after 1990 that have been built in previously unbuilt parcels. It allows estimating the proportion of future housing of the municipality that will be located in previously unbuilt parcels. We obtained:

$$\widehat{HD}_{i,t+s} = \alpha_{1c} \times \widehat{HD}_{i,t+s} \quad \text{with } i \subset c \quad (3)$$

- $\alpha_{2cz}$  is the proportion of housings built after 1990 by type of LUP zones *z* for each type of municipality *c*. It allows

distributing  $\widetilde{HD}_{i,t+s}$  between various LUP zone types of municipality  $i$  depending on the area of the LUP zone  $S_{i,z}$ . Thus:

$$\widetilde{HD}_{i,z,t+s} = \alpha_{3cz} \times \frac{S_{i,z}}{\sum_z S_{i,z}} \widetilde{HD}_{i,t+s} \quad (4)$$

- $\alpha_{3cz}$  is the number of housings by hectare by LUP zone  $z$  and by type of municipality  $c$ .  $H_g(c,z)$  the final number of new houses for the unbuilt cells  $g$  which have a conversion likelihood higher than the local threshold is given by:

$$H_g(c, z) = \alpha_{3cz} \text{ if } \hat{P}(Y_g = 1|X) > \bar{P}(Y_g = 1|X) \quad (5)$$

Moreover,  $\forall g \bar{P}(Y_g = 1|X)$  must verify the two following conditions:

$$\forall i \sum_{g \in i} H_g(c, z) \leq \widetilde{HD}_{i,t+s}$$

$$\forall z \sum_{g \in z \cap i} H_g(c, z) \leq \widetilde{HD}_{i,z,t+s}$$

#### 2.4.4. Model's predictive accuracy

For model (2b) which is a logit model, a simple way to test the model's predictive accuracy can be given by the ratio of true zero and true one. We proposed four model's predictive accuracy and goodness of fit for each county's model based on different samplings. More precisely, the McKelvey and Zavoina pseudo  $R^2$ , and the percentage of "true prediction", of "true one" and of "true zero" are estimated for: (1) the whole sample, (2) a holdout sample 1 (20% randomly excluded), (3) a holdout sample 2 (20% of municipalities randomly excluded), (4) a sample of municipalities of PACA region for which house development over the period 2000–2005 are forecasted with the information set available in 1999.

The proposed measures are based on a threshold value of probability (specific for each zoning/county's model) where the number of false zero equals false one. Indeed, we choose to consider prediction error for development as important as prediction error for non development in our measures of predictive accuracy. Note that the measure of accuracy of prediction for new houses for period 2000–2005 is conducted only in region PACA in which both date of house construction at parcel level and digital cadastral map are available for more than 778 municipalities (80% of PACA municipalities).

The variable simulated by model (1) is continuous and with no contemporary variables in the covariates. Therefore, predictive accuracy can be assessed directly by goodness of fit or RMSE and Theil  $U$  statistics (Greene, 2007). However, such statistics for model (1) are optimistic because it estimates the total number of houses and not the housings growth. Thus, for each sub-regions, we propose the centered  $R^2$  and the Theil  $U$  statistic for model (1) using difference of housings between 1999 and 2006 as explained variable  $HD_{t,i} - HD_{t-1,i}$ . The Theil  $U$  statistic is expressed in that case as  $U_{\Delta} = \sqrt{\sum_i (y_i - \hat{y}_i)^2 / \sum_i (y_i)^2}$  where  $y_i = HD_{t,i} - HD_{t-1,i}$  (or  $y_i = HD_{t,i}$  in the classic case).

### 2.5. Data analysis

#### 2.5.1. Sites considered as threatened

We estimated three urban indicators of pressure at the site level. Pressure 1 (P1) represents the predictable consumption for housing inside each site. P1 is calculated only for the year 2030 and is equal for each site to the percentage that could be newly

built in 2030. A site is considered threatened if  $P1 > 10\%$ . Pressure 2 (P2) corresponds to the smoothed impact of urban area at a local scale in the neighborhood surrounding each site in a 2 km radius. For each cell  $g$ , we calculated P2 as the sum of the number of built-up cells in a 2 km buffer zone weighted by their distance to the cell considered (weights are estimated using a Gaussian kernel smoother with a bandwidth  $h = 0.8$  km). P2 is equal to the mean value of P2 for each site. A site was considered threatened if  $P2 > 85$ 's percentile of all sites in 2008. Pressure 3 (P3) corresponds to the smoothed impact of urban area at a global scale in the large neighborhood of each site. First we calculated P3 as the number of built cells  $g$  in each grid cell of  $1000 \text{ m} \times 1000 \text{ m}$ . For each site we calculated P3 as the mean value of P3 in a 50 km buffer zone weighted by the distance between the 1000 m grid cells and site centroids (weight are estimated using a Gaussian kernel smoother with a bandwidth  $h = 25$  km). Mean values were preferred to totals because they limit border effects due to the large scale distribution of this indicator. A site was considered as threatened if  $P3 > 85$ 's percentile of all sites in 2008. For each indicator of pressure, we calculated the number, mean and total percentage area of threatened sites as well as the mean number of species per threatened site.

An important methodological point of this study is the use of a smoothed function for indicators of pressures 2 and 3. Although several studies have already investigated the indirect impact of human presence on biodiversity surrogates using a buffer zone (Harcourt, Parks, & Woodroffe, 2001; Luck, 2007b; Vazquez & Gaston, 2006), our method allows us to account for a pressure which intuitively decreases in relation to distance from urbanized cells.

#### 2.5.2. Assessing threat levels for species

For each taxonomic group, we investigated the number of threatened species according to the three urban pressures. A species was considered threatened if more than 30% of its range (number of sites) is under pressure. Fish species are not considered for the P1 indicator since they are present in water-courses which are never considered as built-up. We calculated the difference between the number of sites of the species (*i.e.* the range of the species) considered threatened and the others in order to assess the relationship between urban impacts and species rarity. We also calculated for each taxonomic group the mean percentage of threatened sites for each species in order to assess the impact of urbanization beyond the species range threshold. Finally we mapped threatened sites and analyzed the number of sites and species that incur several pressure levels.

## 3. Results

### 3.1. The urban sprawl model

When using LSPR model (1) to forecast housing growth between 1999 and 2006, we found a 0.674 adjusted  $R^2$  (0.9954 for the original model). The  $U$  Theil statistics provides similar results with  $U_{\Delta} = 0.365$  for housings difference and  $U = 0.022$  for total number of housings. Note that housings growth simulated by model (1) has an adjusted  $R^2$  of 0.87 estimated by OLS. The lower goodness of fit for LSPR estimation is the price to pay to have low relative error in small and middle municipalities that are more numerous and more important in the context of future impact of urbanization on ZNIEFF. The loss of predictive accuracy for bigger city is expected to have less impact on urban/ZNIEFF interface due to the spatial location of ZNIEFF that are less present in urban agglomeration. Details

of model (1) results are given in the [Appendix](#) and will be available at a URL after review process.

**Table 2a** gives an overview of goodness of fit and predictive accuracy for each type of zoning  $z$  (URB, FURB, NONDEV) and for the 11 counties as simulated by model (2a). The McKelvey–Zavoina pseudo  $R^2$  varies from 0.68 to 0.86 for URB zoning, between 0.64 and 0.75 for FURB zoning and from 0.38 to 0.69 for NONDEV zoning. The predictive validity measures for sample 1, 2 and 3 provide quite similar results in **Table 2a**. For URB zoning, the minimum percentage of true prediction is around 80% for the less urbanized counties and around 90% for the most urbanized counties (the coastal ones, i.e. for counties numbered 06, 13, 83, 30, 34, 11 and 66). For FURB zoning, the results are more concentrated with approximately 82–85% of true prediction, except for the county 48 with a 97.37% of good prediction. This sub-region is the most rural and non-urbanized and this high value can be explained by a large number of non built cells in FURB zoning. The same high value of true prediction appears for NONBDEV zoning where urbanized cells are rare (between 2 and 9%). When the distribution of 0 and 1 is unbalanced, the percentage of true 1 (or of true 0) gives a better illustration of predictive validity of the model. The percentage of true 1 in NONBDEV varies between 40 and 72%; it is around 60% in coastal counties. In URB zoning where there are less 0, the percentage of true 0 varies between 57 and 74%, with a mean of 70% for coastal counties. The cells forecasted to be developed for the period 2000–2005 in PACA region confirms the predictive accuracy of the models (2a) (see the three last columns of **Table 2a**), notably for NONBDEV where the ratio of true 1 is 3% higher than in the three previous predictive accuracy measures.

In **Table 2b**, in order to highlight the stability of coefficient between counties, we present the value and significance of three coefficients (the closed neighboring urban density  $\beta_{WYp}$ , the distance to small road  $\beta_{dsr1}$ , the dummy “dominant land cover equal meadow in 1990”  $\beta_{CLC90=23}$ ). Results are particularly stable with no sign reversal for the three coefficients (others coefficients present the same regularity, detailed results of the 33 logit models are given in the [Appendix](#) and will be available at a URL after review process). The only exception in **Table 2b** is the effect of the dominant Corine Land Cover (CLC90) that can change among counties. For example, the land cover type meadow ( $\beta_{CLC90=23}$ ) is reversed for the Lozère county (48) which is the less inhabited county of France and where meadows are the dominant land cover. The effect of closed neighboring urban density ( $\beta_{WYp}$ ) is very stable, particularly for zoning URB and FURB. This coefficient decreases generally for most of the inhabited counties (counties 13, 06, 83, 34 and 84 with more than 150 inhabitants/km<sup>2</sup>). We also present and describe in detail the urban model results in **Table 3** only for the *Vaucluse* county (84) and for NONBDEV zoning. The main driving factors influencing the urbanization likelihood (**Table 3**) are location (in suburbs of greatest cities), infrastructure density (i.e. road density within the cell) and housing density in vicinity (i.e. frame of built area in contiguous cells WPy). More precisely, the factors which decrease the urbanization likelihood are slope ( $sl=3$ ) and enforced conservation policy ( $lpa=Y$ ). The covariates which contribute the most to the urbanization likelihood concern the ratio of road inside the cell <10% ( $pr1$ ), the ratio of built surface in contiguous cells (WYp) and to a lesser degree the location in the second commuter belt.

In 2008, 7.5% of the territory of the study was urbanized. Our model projects that additional 1.3% of the territory will be urbanized due to house building by 2030. Regarding Corine Land Cover classes, 30% of new urban areas will occur in already built-up areas, 57% on agricultural land (notably in heterogeneous agricultural zone and permanent culture) and 12% on natural and semi-natural areas (principally in forest and open environment).

**Table 3**

Detailed results of GLM logit model (1) for *Vaucluse* county (84) and for NONBDEV zoning.

Covariates	Coefficients	Pr(> z )
(Intercept)	−1.200e+00	<2e−16***
clc90 reference value = 11		
clc90=12	−1.638e+00	<2e−16***
clc90=13	−2.227e+00	<2e−16***
clc90=14	−2.407e+00	<2e−16***
clc90=21	−1.707e+00	<2e−16***
clc90=22	−1.356e+00	<2e−16***
clc90=23	−1.882e+00	<2e−16***
clc90=24	−1.221e+00	<2e−16***
clc90=31	−2.335e+00	<2e−16***
clc90=32	−2.238e+00	<2e−16***
clc90=33	−2.772e+00	<2e−16***
clc90=41	−1.304e+01	0.94070
clc90=51	−3.849e+00	<2e−16***
sl reference value = 1		
sl=2	−7.186e−01	<2e−16***
sl=3	−8.987e+00	0.77130
zonepop reference value = Urban core		
zonepop=first commuter belt	3.167e−01	1.51e−08***
zonepop=second commuter belt	3.955e−01	2.09e−12***
zonepop=rural area	1.727e−01	0.00621**
nsd reference value = 0		
nsd=1	−4.995e−02	0.02131*
nsd=2	−1.557e−03	0.94260
lpa reference value = N		
lpa=Y	−9.730e−01	0.01061*
duc1	−4.594e−04	<2e−16***
duc2	−3.278e−04	<2e−16***
dsr1	−2.141e−03	<2e−16***
dsr2	−6.011e−04	<2e−16***
pr1	4.817e+00	<2e−16***
pr2	5.856e−01	0.00166**
dmr1	−3.353e−05	<2e−16***
dmr2	−3.160e−05	<2e−16***
area	3.673e−05	1.37e−12***
apr	8.723e−02	0.00296**
WYp	3.094e+01	<2e−16***

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘.’ 1.

Null deviance: 262.406 on 455.791 degrees of freedom.

Residual deviance: 171.764 on 455.760 degrees of freedom.

AIC: 171.828.

### 3.2. Current and future impact on biodiversity

Within the delimited sites, there is an average of 8.5 determinant species per site. The mean surface area of sites is about 823 ha. The mean number of sites per species is about 7.5 for plants, 19.6 for mammals, 21.8 for birds, 18.6 for fishes, 26.8 for herptiles.

The frequency distribution of sites in relation with different pressures in 2030 varies among the three pressure indicators (**Fig. 3a**). The number of threatened species which is assessed from the proportion of their sites being under pressure also varies among the three pressure indicators (**Fig. 3b**). The high number of species which have 100% of their range threatened is likely due to the rare species.

Our model projects that 3512 ha of high diversity sites (i.e. the number of grid cells which will be built) will be impacted by urbanization by 2030. Thus, according to P1, 464 sites (35% of the total number of sites) will have at least 1 ha urbanized between 2006 and 2030. However, only 43 sites will have more than 10% of their surface area urbanized in 2030 (**Table 4**). The surface area of these sites is low (with a mean area of 88 ha) and only 0.4% of the total area of high diversity sites is concerned. Only 36 plant species, one bird and one herptile species could be threatened by direct urbanization. Given the significant low value of the mean number of sites per species for threatened plants (4.7 sites), the



**Table 2a**  
Predictive accuracy of model (2a) in each county and zoning type.

Sub-region	Zoning	(1) Whole sample				(2) Holdout sample 1			(3) Holdout sample 2			(4) 2000–2005 forecasts		
County		McKelvey–Zavoina pseudo $R^2$	% True predict.	%True 0	%True 1	% True predict.	% True 0	% True 1	% True predict.	% True 0	% True 1	% True predict.	% True 0	% True 1
PACA – 04	URB	0.73	79.58	65.25	85.55	79.03	63.92	85.28	78.33	62.79	84.7	80.89	61.82	87.26
	FURB	0.68	82.87	87.86	70.92	82.23	87.47	69.46	90.26	94.07	72.54	78.42	82.27	72.43
	NONBDEV	0.56	98.5	99.24	39.84	98.49	99.23	39.12	98.21	99.09	41.02	96.28	98.08	42.31
PACA – 05	URB	0.68	80.56	57.21	87.42	82.15	59.34	88.59	78.69	54.77	86.06	80.79	53.59	87.89
	FURB	0.64	79.54	83.37	73.41	79.76	83.03	74.93	80.42	84.37	73.8	78.07	81.28	73.54
	NONBDEV	0.65	98.64	99.31	43.09	98.61	99.29	42.04	98.52	99.25	43.95	96.52	98.20	44.22
PACA – 06	URB	0.85	88.43	73.81	92.57	88.31	72.93	92.53	86.79	56.07	92.42	89.59	58.70	94.04
	FURB	0.85	88.25	89.95	85.87	87.12	88.85	84.75	83.24	81.29	84.86	84.63	78.83	87.94
	NONBDEV	0.69	97.77	98.84	72.28	97.8	98.85	72.44	98.28	99.11	71.35	93.12	95.97	76.40
PACA – 13	URB	0.73	91.21	64.73	94.93	91.13	64.73	94.93	91.65	65.59	95.26	92.38	52.30	95.86
	FURB	0.78	85.43	88.76	79.29	85.43	88.71	79.44	87.64	91.64	76.29	80.55	79.02	81.87
	NONBDEV	0.68	93.36	96.39	57.89	93.29	96.36	57.54	92.05	95.58	60.23	85.02	90.75	60.53
PACA – 83	URB	0.80	88.68	72.85	92.85	88.56	72.4	92.79	88.07	72.08	92.38	86.06	63.79	91.36
	FURB	0.75	87.74	90.63	82.24	87.74	90.63	82.28	86.09	89.08	80.86	82.28	84.79	78.77
	NONBDEV	0.59	95.68	97.71	60.12	95.66	97.7	60.26	95.66	97.69	62.91	91.63	95.32	60.61
PACA – 84	URB	0.73	90.02	56.31	94.37	90.21	49.54	96.02	89.51	47.29	95.43	88.00	55.74	93.06
	FURB	0.68	82.75	86.07	77.35	82.79	86.14	77.3	81.44	83.13	79.38	78.63	78.32	78.93
	NONBDEV	0.54	92.3	95.8	54.01	92.2	95.74	54.15	92.05	95.66	52.72	88.00	93.06	55.74
LR – 11	URB	0.86	87.94	74.31	92.11	88.7	75.83	92.62	85.23	90.24	69.61			
	FURB	0.70	84.04	87.44	78.11	83.33	77.34	86.79	83.18	85.98	78.98			
	NONBDEV	0.48	97.77	98.86	44.59	97.76	98.86	44.14	97.83	98.9	42.1			
LR – 30	URB	0.77	88.11	60.23	93.01	87.8	59.58	92.81	86.04	64.3	91.32			
	FURB	0.67	82.23	85.44	77.21	83.32	86.21	78.91	82.74	86.57	75.85			
	NONBDEV	0.49	96.07	97.95	50.83	96.08	97.96	50.76	95.64	97.72	49.89			
LR – 34	URB	0.85	93.5	72.72	96.31	93.49	73.72	96.29	94.48	67.46	97	Data not avail- able		
	FURB	0.69	82.39	81.26	83.4	81.63	80.39	82.73	83.11	81.07	84.75			
	NONBDEV	0.38	97.26	98.59	47.1	97.26	98.6	46.14	97.26	98.6	46.87			
LR – 48	URB	0.75	81.65	71.01	86.58	80.78	68.86	86.14	79.66	67.19	85.26			
	FURB	0.67	97.37	98.59	78.97	97.62	98.73	80.58	96.3	98	75			
	NONBDEV	0.37	98.07	99.01	48.2	98.08	99.02	48.17	98.06	99.01	48.26			
LR – 66	URB	0.82	92.86	64.93	96.02	92.25	60.91	95.7	93.72	63.68	96.6			
	FURB	0.70	82.28	83.68	80.63	81.98	82.69	81.15	82.61	84.89	79.51			
	NONBDEV	0.59	97.98	98.97	52.98	97.99	98.97	53.38	98.48	99.23	52.22			

**Table 2b**

Selected coefficients (the closed neighboring urban density  $\beta_{WYP}$ , the distance to small road  $\beta_{dsr}$  and the dummy “dominant land cover equal meadow in 1990”  $\beta_{CLC90=23}$ ). Estimated results of the GLM logit models (1) for each types of zoning in the 11 counties.

Zoning	County	N	$\beta_{CLC90=23}$	$\beta_{WYP}$	$\beta_{dsr1}$
URB	04	17 427	−0.464***	36.4***	−0.00108***
URB	05	10 331	−0.201 ns	35.7***	−0.000915**
URB	06	29 775	−0.524***	37.1***	−0.00178***
URB	13	48 614	−2.34***	25.3***	−0.0021***
URB	83	45 294	−0.151 ns	38.7***	−0.00135***
URB	84	21 807	−0.625***	25.1***	−0.00101***
URB	11	17 340	−1.22***	33.4***	−0.000296
URB	30	34 577	−1.12***	27.8***	−3.55e−05 ns
URB	34	27 657	−1.42***	30.8***	−0.000679**
URB	48	5384	−0.698***	37.8***	−0.000316 ns
URB	66	12 252	−1.19***	31***	−0.000513*
FURB	04	10 859	−0.516**	35.2***	−0.00273***
FURB	05	4541	−0.724***	34.9***	−0.00244***
FURB	06	13 350	−0.253 ns	28.5***	−0.0035***
FURB	13	49 812	−0.918***	23.4***	−0.00255***
FURB	83	27 331	−0.859***	31.7***	−0.00329***
FURB	84	22 519	−1.8***	25.4***	−0.00145***
FURB	11	14 606	−2.17***	27***	−0.000336*
FURB	30	27 038	−1.84***	25.8***	−0.000704***
FURB	34	30 704	−1.74***	27.3***	−0.000548***
FURB	48	8417	−1.45***	47.6***	−0.000676 ns
FURB	66	13 508	−1.69***	29***	−0.000442**
NONBDEV	04	999 978	−0.431***	63.7***	−0.00299***
NONBDEV	05	679 554	0.0795 ns	66.4***	−0.00319***
NONBDEV	06	536 985	−0.955***	44.7***	−0.00381***
NONBDEV	13	614 993	−2.26***	28.4***	−0.00266***
NONBDEV	83	696 554	−1.22***	52.6***	−0.0027***
NONBDEV	84	455 792	−1.88***	30.9***	−0.00214***
NONBDEV	11	728 446	−1.12***	57.4***	−0.000527***
NONBDEV	30	853 789	−1.11***	40.7***	−0.00137***
NONBDEV	34	700 087	−0.949***	45.1***	−0.00117***
NONBDEV	48	574 364	0.295*	101***	−0.000303***
NONBDEV	66	455 969	−1.19***	29.6***	−0.00123***

ns = non significant.

\*  $p < 0.05$ .

\*\*  $p < 0.01$ .

\*\*\*  $p < 0.001$ .

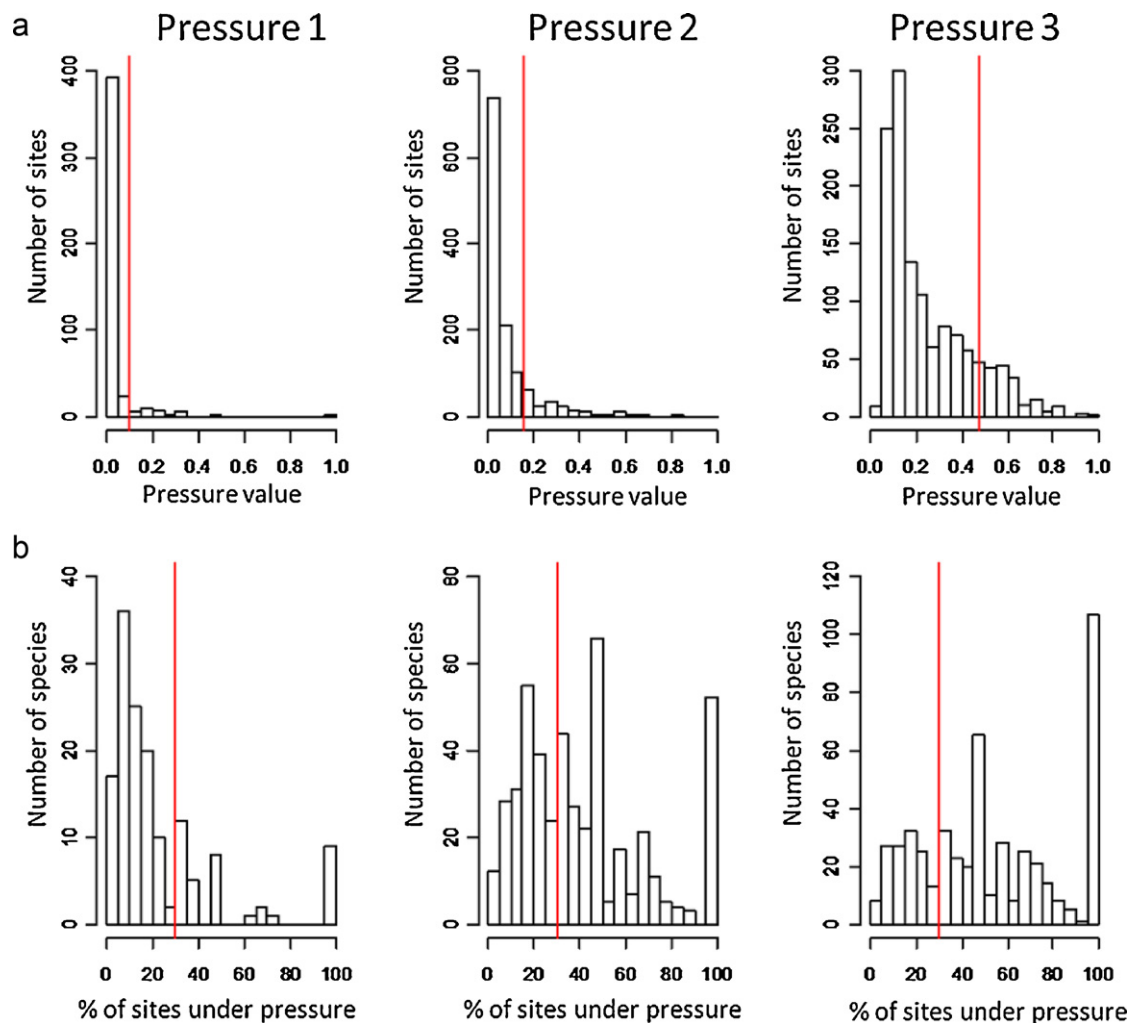
species concerned are rarer than the species which are not threatened (i.e. their mean number of sites is lower). Mammals are the least impacted in terms of mean percentage of threatened sites per species.

According to P2, 198 sites with a mean surface area of 219 ha, representing 4% of the total surface area of sites of conservation interest, are currently threatened by urbanization (Table 4). For each taxonomic group, the species concerned are not rarer than the others. Mammal and herptile species are the least impacted in terms of the mean percentage of threatened sites per species. The most impacted are fish species. In 2030, our model projects an increase in 34% of the number of threatened sites (265 sites). The mean surface area of threatened sites increases from 219 to 309 ha, hence, the surface area of newly threatened sites will be higher in 2030. Based on our thresholds, approximately half (43.8%) of all fish species will be threatened. The fact that the mean number of sites of the threatened species increases for bird species means that the species newly impacted are more widely distributed (in term of number of sites) than those in 2008. In contrast, newly impacted herptile species are rarer than those impacted in 2008. According to P 3, 198 sites, with a mean surface area of 400 ha, are threatened in 2008. This represents 7.3% of the total surface of sites of high conservation interest. No mammal species are threatened and threatened species are not rarer than species unthreatened except for herptile species. In 2030, our model projects an increase in 48% of the number of threatened sites (293 sites). The surface area of newly threatened sites will be higher in 2030. Based on our thresholds, approximately half of

the fish species (47.7%) will be threatened. The mean number of sites of threatened species is similar to that in 2008, except for herptile species which show an increase from 5.3 sites per species in 2008 to 30.1 in 2030. Mammal species are the least impacted in term of the mean percentage of threatened sites per species. Birds and herptiles have the highest percentage of threatened sites.

In 2030, 38 sites of high conservation interest will be threatened by urbanization according to both P1 and P2 (i.e. 83% of the sites threatened under P1 are also threatened under P2). 125 sites will be threatened simultaneously by P2 and P3, and 29 sites will be threatened by P1 and P3. 28 sites are common to all three indicators of pressure 32 plant species and 1 bird species are considered threatened for all three indicators of pressure, while 173 plant species, 10 bird species, 3 fish species and 3 herptile species are considered threatened for at least two of them.

Overall, the majority of sites under pressure for each indicator of pressure are located in the coastal part of the region (Fig. 4). Nonetheless there are some differences regarding the localization of the sites among the three indicators of pressure. Considering P1 and P2, some threatened sites are located in mountain areas (in the north of the study region) while for P3, threatened sites are aggregated around urban poles in the lowland plains, near the coast and in the Rhône valley. Sites newly threatened in 2030 occur close to those already concerned in 2008 for P3 while they are more dispersed for P2.



**Fig. 3.** Frequency distribution in 2030 of (a) the pressure value on sites (when it is >0), (b) the proportion of threatened sites per species (when it is >0) for each of the three pressure indicators. Vertical red lines represent the threshold values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

#### 4. Discussion

Our analysis provides a methodological approach to project future urbanization and to assess the direct and indirect threats on biodiversity at a regional scale. To our knowledge, this is the first study to do so for such large territory and at a so fine grained scale. The three indicators of pressure are highly complementary and range from precise quantification of land consumption to the pressure associated with human presence in a 50 km radius weighted by the distance to the cell. Our study illustrates a diversity of potential impacts of urban spread on sites of high conservation interest and on their associated species in the French Mediterranean region in southern France.

##### 4.1. The urban sprawl model

Our model combines demographic forecast at municipal scale, and spatial forecast of future built location at infra-municipal scale. The main improvement of our model is the spatial localization of housings forecast at cell scale that allows taking into account the location of new built in NONBDEV zoning that may have the most influential threat inside ZNIEFF areas or in their close neighboring (pressures 1 and 2). Even if LUP changes are mainly driven by land market and by land owner anticipation on policy change

(McMillen & McDonald, 1991), the evolution of LUP is generally made by changing rules on a set of cells. In general, contiguous zones with numerous cells with high probability of change constitute the chosen sets. Isolated cells with high probability will likely stay unbuilt because the cells will stay in NONBDEV zoning in the mid-term. At this stage, model (1) takes this phenomenon into account only partially. The spatial dependence of the probability to be built and of zoning changes, would be better accounted for by a Spatial Autoregressive Regression (SAR) logit model which unfortunately cannot be used with such large dataset. An extension of the Klier and McMillen (2008) methodology using generalized GMM for such latent class model may be useful for such rich spatial data context if combined with sparse matrix computation method. The recent work of Diallo and Geniaux (2011) on such extension seems promising. Moreover the development of centralized spatial databases on land use plan, land property and parcel/house polygons in French and European regions guaranty the reproducibility of such modeling framework. It also enables to work at parcel level which is the “decisional scale”: it allows increasing the available covariates of the land conversion decisional process, and will benefits to model adjustment (we have estimated an approximate increase of 20% of the MZ pseudo  $R^2$  comparing cell versus parcel scale for model (2a) for region PACA where parcel data are available).

**Table 4**

Current and future pressures for sites of high conservation interest and the different taxonomic groups for the three pressure levels associated with urbanization.

	2008		2030		
	Pressure 2	Pressure 3	Pressure 1	Pressure 2	Pressure 3
<b>Sites</b>					
Number of sites	198	198	43	265	293
% area of site	4.0	7.3	0.4	7.6	14.8
Mean area of site (ha)	219 <sup>***</sup>	400 <sup>***</sup>	88 <sup>**</sup>	309 <sup>***</sup>	546 <sup>***</sup>
Mean number of species per site	7.2 <sup>*</sup>	8.7 ns	7.2 ns	7.2 <sup>**</sup>	8.5 ns
<b>Plant</b>					
Number of species	192	230	36	252	313
% of species	18.5	22.1	3.5	24.2	30.1
Mean number of sites per species (range)	7.3 ns	7.3 ns	4.7 <sup>***</sup>	7.8 ns	7.9 ns
Mean % of sites under pressure per species	13.1	16.9	2.9	17.0	22.5
<b>Mammal</b>					
Number of species	1	0	0	1	0
% of species	3.6	0.0	0.0	3.6	0.0
Mean number of sites perspecies (range)	–	–	–	9.0	–
Mean % of sites under pressure per species	5.0	2.6	0.2	6.1	5.3
<b>Bird</b>					
Number of species	9	22	1	17	41
% of species	10.5	25.6	1.2	19.8	47.7
Mean number of sites per species (range)	16.3 ns	21.4 ns	–	26.1 ns	24.5 ns
Mean % of sites under pressure per species	11.9	18.4	2.5	16.3	30.2
<b>Fish</b>					
Number of species	6	3	–	7	5
% of species	37.5	18.8	–	43.8	31.3
Mean number of sites per species (range)	8.0 ns	5.3 ns	–	9.4 ns	5.6 ns
Mean % of sites under pressure per species	23.8	15.1	–	30.1	20.1
<b>Herptile</b>					
Number of species	2	3	1	7	8
% of species	10.0	15.0	5.0	35.0	40.0
Mean number of sites per species (range)	41.0 ns	5.3 <sup>**</sup>	–	34.1 ns	30.1 ns
Mean % of sites under pressure per species	9.8	15.9	2.9	18.2	28.4

As explained in Section 2.5.1, pressure 1 corresponds to direct consumption of sites, pressure 2 to the impact of urban area in the nearby neighborhood of each site (2 km), and pressure 3 to the impact of urban area in the large neighborhood of each site (50 km).

ns = non significant.

\*  $p < 0.05$ .

\*\*  $p < 0.01$ .

\*\*\*  $p < 0.001$ .

#### 4.2. Methodological aspects of pressure assessment

Such studies, at regional scale and for a large set of species, require some simplification in the analysis and in the presentation of the results but potential bias must be clearly understood. First, we used some fixed buffer zones in order to assess the indirect pressures 2 and 3 at the site level. In theory this distance is species-specific and should be determined according to the sensitivity of each species considered. Then, the results show mean values at the site level although the population of a particular species will rarely cover systematically an entire site and could be destroyed or become threatened by a localized urban impact in a large site for instance. Finally, we used thresholds in order to identify sites and species under pressure. Even though these thresholds are chosen according to the frequency distribution of the different values (see Fig. 3), each site and each species has different sensitivity to the urban pressure.

One solution to avoid each of these potential biases would be to assess the robustness of the results when changing the different buffer zones, level of aggregation, or thresholds values. Thus, the species or sites of highest priority for conservation would be those determined as under pressure in all the cases.

But, other criteria as site features or species biological traits always need to be taken into account in order to accurately assess the urbanization threats to biodiversity conservation.

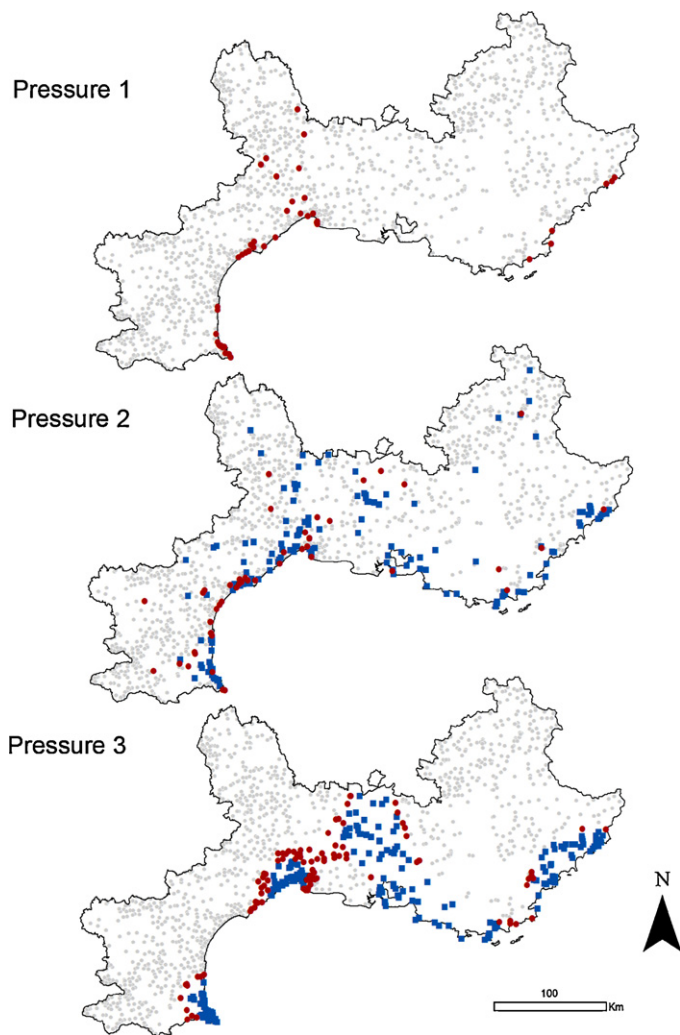
#### 4.3. Current and future impact on biodiversity

Regarding the area of sites and their location as well as, the number of threatened sites and species, the potential impacts of urbanization differ among the three indicators of pressure. Most of the threatened sites of high conservation interest have significantly a smaller area than the other sites and occur primarily in the coastal region of the study area and in the major Rhône river valley. Mammals are the least threatened species by urban pressure. This can be explained by their association with forest habitat in the study region (which is less affected by urbanization than open habitats and agricultural areas in our model) and are present in sites of generally larger area than average.

Direct consumption of sites through urban sprawling is certainly the greatest threat to biodiversity because it leads to an irreversible destruction of habitats and their associated species. Even if the French law indicates that the ZNIEFF inventory of sites of high conservation interest should be accounted for before allowing an area to be urbanized, our study illustrates the high sensitivity of several sites to urban pressure and direct habitat destruction.

Our analysis illustrates that future urbanization could impact as much as 35% of the total number of sites (3512 ha), although less than 50 sites (3.2%) of high conservation interest (having a very small area) will have more than 10% of their area destroyed. It is not surprising that species impacted in more than 30% of sites where they occur are rare plants. Rare plants generally occupy small patch ecosystems with highly localized distributions.





**Fig. 4.** Spatial distribution of threatened sites in relation to three pressure indicators. Blue squares refer to threatened sites in 2008; red points refer to the sites under pressure by 2030; grey points refer to the sites without pressure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

P2 refers to threats induced by immediate proximity of buildings. This indicator of pressure thus depends on the local dynamics and spatial configuration of urbanization and some sites can be under pressure even though they may occur close to only small villages or scattered buildings. It explains why the new sites under pressure in 2030 are not always localized near those already under pressure in 2008. The fishes are the most endangered group of species. It is likely related to the fact that villages and towns are often settled near a river which was used in their development (transport, energy, water resources...). Besides, a narrow site is more likely to be under pressure than a broad site. An increase in 34% of the number of threatened sites in 2030 shows the potential high level of pressure directly around the sites. This pressure illustrates the need to consider one high diversity site in its neighboring environment. Some authors have already discussed the external threats to reserves (DeFries et al., 2007) and suggested the critical importance to create buffer zones around them.

P3 is less sensitive to the local dynamics of urbanization as it provides information on a more regional threat, due to urbanization in a large area of 50 km surrounding the sites. The absolute number of buildings in a global area is likely to be more important than their precise local arrangement. A site located 200 m away from a small village is not necessarily more threatened than a site

located 10 km from a town of 50 000 habitants. Therefore, patterns of variation between the two dates are not the same for P2 and P3. Considering P3, the sites newly impacted in 2030 are aggregated to those already under pressure in 2030. The most impacted sites are those near existing areas of intense urbanization, i.e. in the lowland plains near the coast and in the Rhone valley. Finally, the large-scale impact of this pressure also explains the wider area of sites. Indeed, within a 50 km radius, the size of the site becomes less important to determine the level of impact. At such scale, the indicator of pressure refers to a wide range of indirect urbanization impacts such as road traffic, water and air pollution and is therefore really difficult to predict and to limit. Special attention must be paid site per site in order to assess the potential threat in such sites.

#### 4.4. Implications for conservation strategies

Using indicators of pressure has important implications to determine objective targets and to implement conservation strategies, beyond the insights they provide regarding threat levels associated with direct and indirect urban impacts on biodiversity (Margules & Pressey, 2000; Myers et al., 2000; Pressey & Cowling, 2001). Coupling urban sprawl forecasting model and biodiversity assessment at the French Mediterranean level would be of paramount importance for biodiversity conservation planning. First, it allows identifying the most threatened sites. In our case, these sites are located in the neighborhood of urban agglomerations along the Mediterranean coast and Rhône Valley. Despite, the challenge to maintain biodiversity in coastal urban area is relatively well known (Médail & Diadema, 2006), the Rhône Valley sites were not yet clearly identified as threatened. The model allows anticipating urbanization impacts most likely to occur before urban sprawl limits the possibilities of intervention. Consequently, it could be easier for decision-makers to avoid the isolation of these sites in an urban matrix (preservation of corridors) and manage their frequentation, and even take measures for maintaining traditional land uses. From environmentalists and political decision-makers point of view, the main result of the model could be to point out the relevance of reconciling the preservation of biodiversity and urban sprawl in these territories when it is still rather possible to do it. Moreover, the model is expected to change the scale of analysis of conservation issues related to urban threat by allowing to treat them directly on large regional areas, which is fairly consistent with usual political decision scale. However, this comprehensive approach is not sufficient. Further analyses should take into account not only the main taxonomic groups distribution, but also species diversity and their ecology. Thus, one should have also good knowledge of the situations on the field and a precise characterization of species habitat requirements in order to design conservation policy using such model.

#### Acknowledgments

This work was funded by the Agence Nationale de la Recherche (contract 05-BDIV-014 ABIME), the Ministère de l'Ecologie, de l'Energie, du Développement Durable et de la Mer (contract 0001077-Diva2 BIO2M) and the Languedoc-Roussillon regional government. We are grateful to staff of the Conservatoire des Espaces Naturels and the Conservatoire Botanique National Méditerranéen de Porquerolles for their advice on data.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.landurbplan.2011.11.003](https://doi.org/10.1016/j.landurbplan.2011.11.003).

## References

- Alberti, M. (1999). Modeling the urban ecosystem: A conceptual framework. *Environment and Planning B-Planning & Design*, 26(4), 605–630.
- Araujo, M. B. (2003). The coincidence of people and biodiversity in Europe. *Global Ecology and Biogeography*, 12(1), 5–12.
- Balmford, A., Moore, J. L., Brooks, T., Burgess, N., Hansen, L. A., Williams, P., et al. (2001). Conservation conflicts across Africa. *Science*, 291(5513), 2616–2619.
- Batty, M. (1991). Generating urban forms from diffusive growth. *Environment and Planning A*, 23(4), 511–544.
- Benoit, G., & Comeau, A. (2005). (*Mediterranea. The plan bleu perspectives on environment and development*) Méditerranée. Les perspectives du Plan Bleu sur l'environnement et le développement. Editions de l'Aube. (in French)
- Bessy-Pietri, P. (2000). Les formes récentes de la croissance urbaine [The recent shapes of urban growth]. *Economie et statistique*, 336(6), 35–52 (in French).
- Blondel, J., Aronson, J., Bodiou, J.-Y., & Boeuf, G. (2010). *The Mediterranean basin – Biological diversity in space and time*. Oxford, UK: Oxford University Press.
- Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., et al. (2000). Consequences of changing biodiversity. *Nature*, 405(6783), 234–242.
- Clergeau, P., Savard, J. P. L., Mennechez, G., & Falardeau, G. (1998). Bird abundance and diversity along an urban-rural gradient: A comparative study between two cities on different continents. *Condor*, 100(3), 413–425.
- Conway, T. M., & Lathrop, R. G. (2005). Modeling the ecological consequences of land-use policies in an urbanizing region. *Environmental Management*, 35(3), 278–291.
- DeFries, R., Hansen, A., Turner, B. L., Reid, R., & Liu, J. G. (2007). Land use change around protected areas: Management to balance human needs and ecological function. *Ecological Applications*, 17(4), 1031–1038.
- Diallo, A., & Geniaux, G. (2011). Spatial logit for large samples using linearized GMM estimator. In *V World conference of the spatial econometrics association* (Available on demand).
- Fagan, W. F., Meir, E., Carroll, S. S., & Wu, J. G. (2001). The ecology of urban landscapes: Modeling housing starts as a density-dependent colonization process. *Landscape Ecology*, 16(1), 33–39.
- Fontaine, C. M., & Rounsevell, M. D. A. (2009). An agent-based approach to model future residential pressure on a regional landscape. *Landscape Ecology*, 24(9), 1237–1254.
- Forman, R. T. T. (2008). The urban region: Natural systems in our place, our nourishment, our home range, our future. *Landscape Ecology*, 23(3), 251–253.
- Geniaux, G. (2010). *URBANSIMUL 1. Un modèle prototype de simulation de l'occupation du sol à l'échelle parcellaire*. Paris: RICS COBRA. (in French, Available on demand).
- Geniaux, G., Ay, J.-S., & Napoléone, C. (2011). A spatial hedonic approach on land use change anticipations. *Journal of Regional Science*, 51(5), 967–986.
- Geniaux, G., Dumas, E., Bartoli, C., Cezanne-Bert, P., & Napoléone, N. (2005). *Identification qualitative des espaces disponibles pour l'urbanisation nouvelle [Qualitative assessment of the available space for urbanization]*. Rapport pour le Conseil régional Provence Alpes Côte d'Azur/Association CESSA (in French).
- Geniaux, G., & Napoleone, N. (2008). Semi-parametric tools for spatial hedonic models: An introduction to Mixed Geographically Weighted Regression and Geoadditive models. In *Hedonic methods in housing markets – Pricing environmental amenities and segregation*. New York: Springer.
- Geniaux, G., Podjeski, C., & Leroux, B. (2009). Les données MAJIC et leur valorisation au service de l'observation foncière [The Majic data and their valorisation for land observation statistic]. *Etudes Foncières*, 139, 28–32 (in French).
- Greene, W. (2007). *Econometric analysis* (newest edition). Prentice-Hall.
- Grimm, N. B., Grove, J. M., Pickett, S. T. A., & Redman, C. L. (2000). Integrated approaches to long-term studies of urban ecological systems. *Bioscience*, 50(7), 571–584.
- Harcourt, A. H., Parks, S. A., & Woodroffe, R. (2001). Human density as an influence on species/area relationships: Double jeopardy for small African reserves? *Biodiversity and Conservation*, 10(6), 1011–1026.
- Hastie, T. J., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society B*, 55(4), 757–796.
- Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. Chapman & Hall/CRC.
- INPN. (2006). *L'inventaire Znieff*. Paris: INPN.
- Irwin, E. G. (2010). New directions for urban economic models of land use change: Incorporating spatial dynamics and heterogeneity. *Journal of Regional Science*, 50(1), 65–91.
- Irwin, E. G., & Geoghegan, J. (2001). Theory, data, methods: Developing spatially explicit economic models of land use change. *Agriculture Ecosystems & Environment*, 85(1–3), 7–23.
- Jarrige, F. (2004). Les mutations d'une agriculture méditerranéenne face à la croissance urbaine. Dynamique et enjeux autour de Montpellier [Mediterranean mutations in agriculture in urban development context. Dynamic and stakes around Montpellier]. *Cahiers Agricultures*, 13, 64–74 (in French).
- Jenerette, G. D., & Potere, D. (2010). Global analysis and simulation of land-use change associated with urbanization. *Landscape Ecology*, 25(5), 657–670.
- Klier, T., & McMillen, D. (2008). Clustering of auto supplier plants in the United States. *Journal of Business and Economic Statistics*, 26(4), 460–471.
- Lambin, E. F. (1997). Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography*, 21(3), 375–393.
- Landis, J. D. (1995). Imagining land-use futures – Applying the California urban futures model. *Journal of the American Planning Association*, 61(4), 438–457.
- Lawson, D. M., Lamar, C. K., & Schwartz, M. W. (2008). Quantifying plant population persistence in human-dominated landscapes. *Conservation Biology*, 22, 922–928.
- Luck, G. W. (2007b). The relationships between net primary productivity, human population density and species conservation. *Journal of Biogeography*, 34(2), 201–212.
- Luck, G. W. (2007a). A review of the relationships between human population density and biodiversity. *Biological Reviews*, 82, 607–645.
- Luck, M. A., Jenerette, G. D., Wu, J. G., & Grimm, N. B. (2001). The urban funnel model and the spatially heterogeneous ecological footprint. *Ecosystems*, 4(8), 782–796.
- Manley, P. N., Parks, S. A., Campbell, L. A., & Schlesinger, M. D. (2009). Modeling urban land development as a continuum to address fine-grained habitat heterogeneity. *Landscape Urban Planning*, 89(1–2), 28–36.
- Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, 405(6783), 243–253.
- Martinuzzi, S., Gould, W. A., & Gonzalez, O. M. R. (2007). Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. *Landscape Urban Planning*, 79(3–4), 288–297.
- McDonald, R. I., Kareiva, P., & Forman, R. T. T. (2008). The implications of current and future urbanization for global protected areas and biodiversity conservation. *Biological Conservation*, 141(6), 1695–1703.
- McKinney, M. L. (2002). Urbanization, biodiversity, and conservation. *Bioscience*, 52(10), 883–890.
- McMillen, D. P., & McDonald, J. F. (1991). Urban land value functions with endogenous zoning. *Journal of Urban Economics*, 29(1), 14–27.
- Medail, F., & Quezel, P. (1999). Biodiversity hotspots in the Mediterranean basin: Setting global conservation priorities. *Conservation Biology*, 13(6), 1510–1513.
- Médail, F., & Diadema, K. (2006). Biodiversité végétale méditerranéenne et anthropisation: Approches macro et micro régionales. *Annales de Géographie*, 651, 618–640.
- Meyer, W. B., & Turner, B. L. (1992). Human-population growth and global land-use cover change. *Annual Review of Ecology and Systematics*, 23, 39–61.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858.
- Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers*, 93(2), 314–337.
- Pogodzinski, J., & Sass, T. (1994). The theory and estimation of endogenous zoning. *Regional Science and Urban Economics*, 24(5), 601–630.
- Pressey, R. L., & Cowling, R. M. (2001). Reserve selection algorithms and the real world. *Conservation Biology*, 15(1), 275–277.
- Rouget, M., Richardson, D. M., Cowling, R. M., Lloyd, J. W., & Lombard, A. T. (2003). Current patterns of habitat transformation and future threats to biodiversity in terrestrial ecosystems of the Cape Floristic Region, South Africa. *Biological Conservation*, 112(1–2), 63–85.
- Shi, H., Singh, A., Kant, S., Zhu, Z. L., & Waller, E. (2005). Integrating habitat status, human population pressure, and protection status into biodiversity conservation priority setting. *Conservation Biology*, 19(4), 1273–1285.
- Taylor, J. J., Brown, D. G., & Larsen, L. (2007). Preserving natural features: A GIS-based evaluation of a local open-space ordinance. *Landscape Urban Planning*, 82, 1–16.
- Theobald, D. M., & Romme, W. H. (2007). Expansion of the US wildland-urban interface. *Landscape Urban Planning*, 83(4), 340–354.
- Thompson, J. D. (2005). *Plant evolution in the Mediterranean*. Oxford: Oxford University Press.
- Thorson, J. (1994). Zoning policy changes and the urban fringe land market. *Real Estate Economics*, 22(3), 527–538.
- Tofallis, C. (2009). Least squares percentage regression. *Journal of Modern Applied Statistical Methods*.
- Vazquez, L. B., & Gaston, K. J. (2006). People and mammals in Mexico: Conservation conflicts at a national scale. *Biodiversity and Conservation*, 15(8), 2397–2414.
- Veldkamp, A., & Lambin, E. F. (2001). Predicting land-use change. *Agriculture Ecosystems & Environment*, 85(1–3), 1–6.
- Verbarg, P. H., Schot, P. P., Dijst, M. J., & Veldkamp, A. (2004). Land use change modelling: Current practice and research priorities. *GeoJournal*, 61, 309–324.
- Wallace, N. (1988). The market effects of zoning undeveloped land: Does zoning follow the market? *Journal of Urban Economics*, 23(3), 307–326.
- Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A., & Losos, E. (1998). Quantifying threats to imperiled species in the United States. *Bioscience*, 48(8), 607–615.
- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC.
- Wu, X., Hu, Y., He, H., Bu, R., & Xi, F. (2007). Spatiotemporal pattern and its driving forces of urban growth in Shenyang City. *Yingyong Shengtai Xuebao*, 18(10), 2282–2288.