

Towards an Environmental DSS based on Spatio-Temporal Markov Chain Approximation

G rard Balent^a, Marc Deconchat^a, Sylvie Ladet^a, Roger Martin-Clouaire^b and R gis Sabbadin^b

^aINRA-DYNAFOR, BP 27, 31326 Castanet-Tolosan Cedex, France

^bINRA-BIA, BP 27, 31326 Castanet-Tolosan Cedex, France

Abstract: The aim of this paper is to provide a mathematical model for assessing the influence of forest fragmentation on the dynamics of animal biodiversity in a changing landscape. The model is based on a stochastic, spatially explicit population dynamics model which takes both temporal and spatial dynamics of biological processes into account. Unfortunately, this model is not tractable, so we will use a Monte Carlo simulation method in order to approximate the multidimensional random variables involved.

The main strength of our approach is its ability to model generic biological and socio-economic dynamic processes, which are both explicitly spatial and stochastic. In order to demonstrate the usefulness of our biodiversity dynamics modeling tool we use available spatial data on the presence/absence of *Erithacus Rubecula* (robin) at different time points in the “Vall e de la N re”, an area of fragmented forest located in the southwest of France, near Toulouse.

Keywords: Landscape fragmentation, population dynamics, spatio-temporal Markov Chain approximation.

1 INTRODUCTION

Conservation and biodiversity management are important issues, especially in places where global climatic or landscape changes (fragmentation) may drastically transform the ecosystem, with positive or negative influences upon human activities Huston [1994]. Understanding and anticipating these changes requires assessment of large regions in a quick and reliable way, but most predictive models of biodiversity operate at fine-grained spatial scale Deconchat and Balent [2001], or require a great amount of information Conroy and Noon [1996]. Remote sensed data provide a unique way to obtain habitat description over large areas, provided that less precise prediction is accepted Williams [1996]. The main difficulty is to establish a good statistical relationship between a set of species occurrence observations and the data sensed from space.

Such a model has been proposed by Lauga and Joachim [1992], which can be applied over large areas to produce a map of presence probabilities for a given species. However, this “static” approach neither takes into account the dynamics of the processes involved, nor the uncertainty pervading them.

The aim of this paper is to tackle these aspects by providing a model of the influence of forest fragmentation on the dynamics of animal biodiversity in a changing landscape. More precisely, we provide a mathematical model for studying the effect of landscape use change on biodiversity. The main strength of our approach is its ability to model generic biological and socio-economic dynamic processes, which are both explicitly spatial and stochastic.

Our method will be illustrated by a study of the dynamics of robins (*Erithacus Rubecula*) in a fragmented area of the southwest of France. However, our point is not to contribute to the knowledge of robin’s biology, but rather to propose a general-purpose modeling tool. This is why we will just use well-known data on robin’s biology and an empirical study of the “static” response of the bird’s presence to the forest index. The missing parameters of the dynamic model (mainly dispersion) will be given “plausible” values or will be adjusted so that the long-term probabilities of presence of birds computed by our dynamic model will converge to the ones computed by an existing “static” response model.

In other words, given that the only data that we have at the present time were collected in order to build a response of robin's presence to an index called Forest Index (FI) Ladet [2000], we will consider that the output of this static model can be seen as the set of equilibrium (long-term) robin presence probabilities. So, when calibrating our dynamic model we will try to make them converge in the long run to the output of the static model. Of course, it should be precised that the interest of our model is not in its ability to mimic the output of the static model, but its ability to model short-term, out-of-equilibrium changes in the biological processes, in response to expected socio-economic changes! However, we have not sufficient data yet for assessing the quality of our tool with respect to observed changes in landscape-use and presence probabilities. These are difficult to acquire since they need following individual for several years, on a significantly large area. However, we hope that our modeling tool we help to i) give some general ideas on the impact of socio-economic expected changes (forest fragmentation, agricultural land desertion...) on birds presence and ii) help focusing costly data collection by underlying which parts of the biological model are important for assessing such an impact, by using sensitivity analysis on the range of plausible values of parameters, for example.

We will present the *static* Forest Index response model in the following section. Then, we will point out some limitations of the static model, and present an improvement in Section 3, consisting in modeling the dynamics of robin through the use of a Markov chain on a multidimensional random variable, approximated by a set of pseudo independent mono-dimensional random variables. The obtained model is finally tuned and validated through comparisons with the static model and with on field measures (Sections 4 and 5).

2 STATIC MODEL OF PRESENCE / ABSENCE OF ROBIN

The study area lies between the Garonne and Gers rivers, in South-western France (lat.: N43°, long.: W1°). It is a hilly region (200-400m a.s.l.), dissected by north-south valleys, within a sub-Atlantic climate with Mediterranean and mountain influences. The forests are fragmented and cover 15% of the area. Oaks *Quercus robur* and *Q. sessiflora*, often in association with chestnut *Castanea sativa* in coppice, cherry *Prunus avium* and wild service trees *Sorbus torminalis* are the main tree species in the area. The avian fauna is rather poor, and typical

of French plain regions Joachim et al. [1997]. For each plot, experienced observers recorded all bird species contacted visually or by their vocal manifestation during 20 minute periods between sunrise and up to 4 hours after sunrise. The bird census was performed during the month of May 1990 and included 676 points scattered over the area. For the present study, we retained only the presence/absence information of robin. The SPOT satellite images cover a region of 60 x 60 km centered on N43° latitude and E1° longitude. The picture has been windowed on a study zone of approximately 2100 km², with a 20m resolution. As we know that Robin is strongly influenced by forest density and fragmentation, we classified the images with supervision to produce a binary map (forest/not forest).

According to previous works in the region Lauga and Joachim [1992], Ladet [2000], we compute for each points of the map an index of forest influence (FI). The FI of a given point lies between 0 and 1, 0 in an open area and 1 in a completely forested area. In order to compute the FI in a cell of coordinates (i, j) , we take into account the presence or absence of forested cells within a radius r around the cell. Furthermore, the influence of forested cells is smaller when cells are further away. So, cells are weighted according to their distance to the cell in which the FI is computed, the weight decreasing with the distance. In the case of robins, the value of the radius has been set to 100m Lauga and Joachim [1992]: cells further than 100m from the considered cell do not influence the FI. Let now B be the binary matrix of forested and non forested cells (resolution of 20m). Let $W(h) = \exp(\text{param} * h)$ where $\text{param} = \log(0.1)/(r - 1)$ be the (decreasing) weight of distant cells.

Then if a given cell has coordinates (i, j) ,

$$FI(i, j) = \sum_{(i', j'), (i-i')^2 + (j-j')^2 \leq r^2} B(i', j') \times W(i', j').$$

The forest influence combines information on both forest patch size and isolation. These two variables have effects on birds occurrence Lescourret and Genard [1993], Villard et al. [1999]. A logistic regression linked the FI values and the presence/absence of Robin measured on the sampling points. The maximum of likelihood estimated the quality of the model. We cannot determine a priori the range of forest influence on Robin: despite it is a small species, we know that long-range influences can occur, because of a source effect of large forests for example Monteil et al. [2004]. In order to find the best model of occurrence we tested several radiuses for FI and produced a logistic model

for each of them. Then, we chose the radius providing the model with the lowest maximum of likelihood (100m). We applied this model of occurrence on the whole studied area to predict robin's distribution. The frequency of presence increased with the forest influence (FI), allowing using a simple logistic regression for modeling its response. Indeed robin is a bird of forest interior and forest edge in the area under study.

3 SPATIALLY EXPLICIT STOCHASTIC DYNAMICS MODEL

The problem with the robin's presence model we have just presented is twofold: i) it is purely static and cannot take any colonization effect into account and can hardly model the short-term impact of clear-cutting highly densely populated forest areas and ii) it is not really "spatial" insofar as, for example, two equally fragmented areas will have the same probability of presence of robin, even if they are located in different parts of the forest. In this Section we try to remedy these problems by coupling with the FI model a stochastic, individual based, spatially explicit population dynamics model in order to take into account the temporal and spatial dynamics of robin. Unfortunately, this model is not tractable, due to the high dimensionality of the random variables involved. So, in Section 4 we will use a Monte Carlo simulation method in order to approximate the probabilities of presence of robin across the area under study.

3.1 Robin's presence in the area

The area under study is represented by an array of $N \times M$ cells, each of which representing a surface of $100m \times 100m$. The cell surface is chosen with respect to the usual individual territory (1ha), since robin is a territorial bird and there are in general no more than one nest in each cell. So, 20m resolution cells are grouped by 25 (5×5) and an average FI is computed for each group. X_t of dimensions $N \times M$ is the random variable representing the spatial configuration of robin's nests over the whole territory. $p_t(i, j)$ is a variable stating whether a nest is present in cell (i, j) . $p_t(i, j) = 1$ if there is a nest in cell (i, j) at time step t (i.e. year t) and $p_t(i, j) = 0$ if not. So the state space of X_t is $\{0, 1\}^{N \times M}$. We make the assumption that X_t is an homogeneous Markov chain: $\forall t, \forall p_0, \dots, p_{t+1} \in \{0, 1\}^{N \times M}$, $P(X_{t+1} = p_{t+1} | X_0 = p_0, \dots, X_t = p_t) = P(X_{t+1} = p_{t+1} | X_t = p_t)$ So X_t only depends on X_{t-1} , i.e. on the spatial configuration of the nests the previous year.

In what follows, we approximate X_t by a product of "independent" mono-dimensional random variables $X_t^{i,j}$ where $X_t^{i,j}(1) = P(p_t(i, j) = 1)$. X_t will thus be a kind of map of indices of presence of birds on the territory. We will be interested in computing the evolution of the map X_t over time. Since computing this map needs computing the $M \times N$ dimensional random variable X_t , it is generally impossible with large maps. This is the reason for the drastic simplification we use, which will lead to a crude approximation of the exact map X_t .

3.2 Population dynamics model

Clearly, we do not aim at studying robin's biology, but only to illustrate our method on the robin's example. This is why we do not expect the following data to be accurate or even to sound completely reasonable to a biologist. We just collected some data from the literature on the biologic traits of robins Isenmann [2003], and for the unknown parameters we tuned them so that the long term limit of the distribution fits the static model. This is only to be seen as an illustration and any real application of our work would need a huge bibliographic and experimental effort to get the real values of parameters.

We took as parameters :

- Robin's clutch have four to six eggs.
- They have one or two clutches a year.
- Females start reproducing at the age of one.
- The fledging rate is approximately 45%.
- 60% of the birds die every year.

In addition, in general robins build new nests within a distance of 10 km from their parents'. We used these scarce data in order to build a spatially explicit population dynamics model. The model we built is largely arbitrary, due to the poor data we have. However, we should give a few remarks : i) Our aim is to prove the feasibility of a spatially explicit population dynamics model on a large state space, and not in a first stage to be as exact as possible, ii) unknown parameters can be tuned in order to fit the measures on the field, iii) one advantage of the model is to point out the parts of it which need additional experimentations in order to be improved.

Reproduction. From the data, we get that females have in average ten eggs each year. Since the fledging rate is around 45%, there remain around four viable eggs and thus (we assume the sex ratio to be around 0.5), two females. Since the average yearly mortality is 60% and the sexual maturity is attained at one year of age, the two remaining fe-

male have only 40% of chance of generating a new nest. Taking these data into account we chose to fix the following probabilities on the number of successor nests for each nest: $P_{succ}(0) = 0.6$; $P_{succ}(1) = 0.06$; $P_{succ}(2) = 0.14$; $P_{succ}(3) = 0.14$; $P_{succ}(5) = 0.06$. These probabilities are arbitrary but can be adjusted by on site studies.

Mortality. We first assumed a fixed mortality rate of 60%. The figures above give a fixed expectation of the number of successors for any given nest. Namely, $E[Succ] = 0.6 \times 0 + 0.06 \times 1 + \dots + 0.06 \times 4 = 1$. So, the expected number of nests may stay constant when we neglect any effect of the diffusion. If we had chosen $E[Succ] < 1$, the population would gradually decline to 0, while if $E[Succ] > 1$ after a while the whole territory would be populated... These three possible evolutions do not fit the reality in which, when the landscape is modified in one way or another, the nests population varies until it reaches a fixed point close to the FI response. This is why we chose to vary the mortality rate as a function of the forest index FI: Nests that are located in open areas are more subject to predation (since they are more visible) than nests located in completely forested areas. So, we define the mortality rate $\mu(i, j)$ for the cell of coordinates (i, j) as:

$$\mu(i, j) = a \cdot \frac{e^{b_0 + b_1 \cdot FI(i, j)}}{1 + e^{b_0 + b_1 \cdot FI(i, j)}} + b \quad (1)$$

This is the definition of a logistic response of mortality to forest index. a , b , b_0 and b_1 are parameters which will be tuned in Section 5

Diffusion. It is likely that the parameters which influence the location of the successor nests of a given nest are **the distance from the parent's nest**, **the forest index in the area** and **the presence of other nests**. In the absence of information on actual diffusion of nests, we chose to model the diffusion by a simple 2D truncated Gaussian probability function of the distance to the parent nest. The probability of a nest being located at a distance d from its parent nest is: 0 if $d > 5km$, and the value of the Gaussian of average 0 and 99% quantile of 5km if $d \leq 5km$. This simple definition of the diffusion does not take into account the forest index of the target cell. In reality birds are most likely to build a new nest in a favorable environment. This is why we chose to weight the Gaussian distribution by an increasing function of the forest index. Indeed, the new probability that a child nest be located in cell (i, j) given that its parent nest is located

in cell (i_0, j_0) is:

$$P_{dif}(i, j, i_0, j_0) = \alpha \cdot G(i, j, i_0, j_0) \cdot W(i, j) \\ \text{where } W(i, j) = \frac{e^{c_0 + c_1 \cdot FI(i, j)}}{1 + e^{c_0 + c_1 \cdot FI(i, j)}} + d \quad (2)$$

where α is a normalization factor, c_0 , c_1 and d are parameters and $G(i, j, i_0, j_0)$ is the 2D Gaussian of parameters $(i - i_0, j - j_0)$ computed above. Finally, in order to take the presence of other nests into account we modified further the probability of diffusion by making it 0 in cells where a nest already exists.

So, we now have a stochastic, spatially explicit model of robin's nests dynamics. With this model we could imagine to handle the $N \times M$ -dimensional random variable X_t representing the possible spatial configurations of nests in the territory. However, this is impossible in practice, due to the high dimensionality of X_t . This is why we introduce in the following section an approximation of X_t by a set of "independent" probabilities of presence $X_t^{i, j}$, which evolution over time will be computed through simulation Sabbadin [2003].

4 PROBABILITIES OF PRESENCE APPROXIMATION

We propose to approximate the multidimensional random variable X_t by a set of pseudo-independent random variables $\{X_t^{i, j}\}_{(i, j) \in N \times M}$. It is only an approximation: it can be shown easily that X_{t+1} is not a product of independent variables. We decided to compute these pseudo-independent distributions by Monte Carlo simulation.

4.1 Monte Carlo simulation of the evolution of the probabilities of presence

$\{X_t^{i, j}\}_{i=1 \dots N, j=1 \dots M}$ being given, i.e. assuming that at time step t we know the probabilities of presence of nests and assuming (here is the approximation) that these are independent, we suggest to compute $\{X_{t+1}^{i, j}\}_{i=1 \dots N, j=1 \dots M}$ by simulation, using the following algorithm. We will repeat for an arbitrarily large number of iterations `nbiter`:

1. Draw $(p_t(i, j))_{i=1 \dots N, j=1 \dots M}$ from $\{X_t^{i, j}\}_{i=1 \dots N, j=1 \dots M}$.
2. For each (i, j) such that $p_t(i, j) = 1$, Simulate the mortality, reproduction and migration from cell (i, j) . The simulated population is denoted $(y_t^{i, j}(i', j'))_{i'=1 \dots N, j'=1 \dots M}$.
3. Let now $n_{t+1}(i', j') \leftarrow n_t(i', j') + \sum_{i=1 \dots N, j=1 \dots M} (y_t^{i, j}(i', j'))$ (where

$(n_{t+1}(i, j))_{i=1\dots N, j=1\dots M}$ is initialized to 0 at the beginning of the algorithm).

Now, after `nbiter` iterations, we have an array $(n_{t+1}(i, j))_{i=1\dots N, j=1\dots M}$ of numbers between 0 and `nbiter`. The simulated random variable $\{X_{t+1}^{i,j}\}_{i=1\dots N, j=1\dots M}$ is simply given by $X_{t+1}^{i,j}(1) = n_{t+1}(i, j)/\text{nbiter}$ for all i, j .

4.2 Comparison with the FI response

We have proposed a means of computing the evolution of the maps of probabilities $\{X_t^{i,j}\}_{i=1\dots N, j=1\dots M, t=0\dots+\infty}$. In order to relate this way to study the dynamics of the robin with the static Forest Index response method, we can compare on the Vallée de la Nère example the FI response index and the $(X_{+\infty}^{i,j})$ map. They should be at least qualitatively similar. We are not looking for an exact similarity since we already pointed out the limitations of the FI response. However, as Figures 2 and 3 show, the two maps are similar. Note that the $(X_{+\infty}^{i,j})$ map was obtained after tuning the parameters of the dynamic model (the tuning process is explained in section 5). However, we can point out a qualitative difference between the FI response and the dynamic model. Namely, the probability of presence of robins in small isolated forests, computed by the dynamic model, is smaller than the response to the Forest Index. On the contrary, the probability of presence of robins in large forests is slightly overestimated.

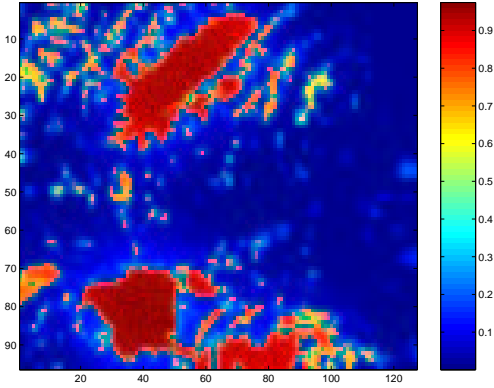


Figure 1: Limit probabilities of presence.

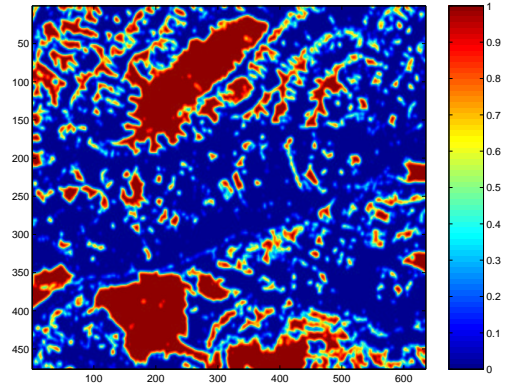


Figure 2: Forest Influence response

5 TUNING OF THE PARAMETERS AND VALIDATION OF THE DYNAMIC MODEL

5.1 Dynamic model versus FI response

Forest Index response. The formula that gives the probability of presence as a function of the Forest index is the following Ladet [2000]:

$$P(FI) = \frac{e^{B_0 + B_1 \cdot FI}}{1 + e^{B_0 + B_1 \cdot FI}} \quad (3)$$

With parameters $B_0 = -1.987$ and $B_1 = 4.891$.

Dynamic model. The *reproduction*, *mortality* and *diffusion* parameters can be adjusted. We chose not to modify the reproduction parameters. So, for the two other parameters: i) *Mortality*. We chose to model the influence of FI on mortality by making mortality a function of FI as shown in Equation 1. ii) *Diffusion*. Diffusion parameters $P_{dif}(i, j, i_0, j_0)$ where also functions of FI as shown in Equation 2. In order to tune parameters a, b, B_0, B_1, C_0, C_1 and d , we compared the limit map X_∞ with the FI response map. To be more precise, we chose an initial distribution X_0 identical to the FI response ($X_0^{i,j} = r(FI(i, j))$, where $r(FI(i, j))$ is the response to the FI of cell (i, j)). Then, our dynamic model should stay close to the FI response over time, thus ensuring that the FI response map is a limit distribution of the process. In order to check this, we simulated the dynamics of X_t until convergence, and then compared the cloud of points representing the limit probabilities of presence observed as functions of the FI, with the logistic FI response. Several comparisons were made and the best fit is shown on Figure 4.

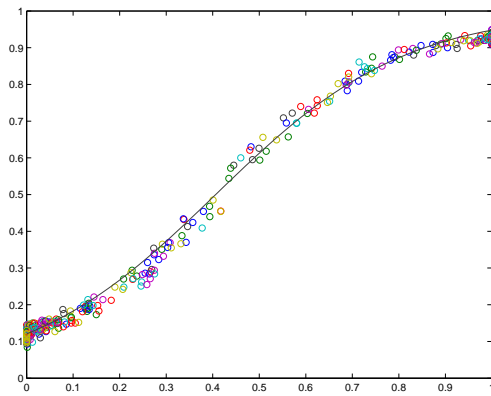


Figure 3: Limit presence probabilities as functions of FI for $B_0 = -3$, $B_1 = 4.891$, $a = -0.2$, $b = 0.93$, $C_0 = -5$, $C_1 = 6$ and $d = 0.007$

6 CONCLUDING REMARKS

In this paper we have provided a model which can be used for the study of spatio-temporal dynamics of biological processes in spatial, possibly dynamic landscapes. This model explicitly handles spatial features, as well as stochasticity of events. The main originality of the model is to use a crude approximation of a multidimensional random variable, which evolution (otherwise impossible to model) over time is estimated through Monte Carlo Simulation. It has been illustrated on the example of robin in a large valley of the southwest of France.

This work is only preliminary and we are considering the following extensions:

- i) First, the use of the model on the example of robins need further experiment to be validated.
- ii) Other birds species with different habitat requirements have been studied with respect to their relation to FI (44 species) in Ladet [2000]. The dynamics model we propose here should be adapted to these species as well, in order to have a clear view of the biodiversity evolution in the area.
- iii) In the dynamic model, only animal dynamics is considered, not landscape evolution. However, the forest cover of the area under study has known deep changes in the past and should encounter even more changes in the near future. It is clear that our objective is to measure the impact of such changes on biodiversity. This could be easily studied within our framework, by considering landscape dynamics in addition to animal dynamics. In this way, the model could help decision makers in assessing the impact of large scale decisions (deforestation, land consolidation) on biodiversity.

REFERENCES

- M. Conroy and B. Noon. Mapping of species richness for conservation of biological diversity: conceptual and methodological issues. *Ecological Applications*, 6:763–773, 1996.
- M. Deconchat and G. Balent. Vegetation and bird community dynamics in fragmented coppice forests. *Forestry*, 74:105–118, 2001.
- M. Huston. *Biological diversity. The coexistence of species on changing landscapes*. Cambridge University Press, Cambridge, UK, 1994.
- P. Isenmann. *Le rougegorge*. Belin-Eveil Nature, Paris, 2003.
- J. Joachim, J. Bousquet, and C. Faure. *Atlas des oiseaux nicheurs de Midi-Pyrénées. Années 1985 à 1989*. AROMP, Toulouse, France, 1997.
- S. Ladet. Modélisation de la réponse des espèces d’oiseaux à la fragmentation forestière dans les coteaux du sud-ouest. Master’s thesis, Université Paul Sabatier. DEA Ecologie des systèmes continentaux, 2000.
- J. Lauga and J. Joachim. Modelling the effects of forest fragmentation on certain species of forest-breeding birds. *Landscape Ecology*, 6:183–193, 1992.
- F. Lescourret and M. Genard. Habitat, landscape and birds composition in mountain forest fragments. *Journal of Environmental Management*, 40:317–328, 1993.
- C. Monteil, M. Deconchat, and G. Balent. Interest of neural network for landscape ecology- example of bird species richness models in fragmented forests. *Landscape Ecology*, 2004.
- R. Sabbadin. Approximating spatial markov decision processes for environmental management. In *15th International Congress on Modelling and Simulation (MODSIM’03)*, volume vol. 4, pages 1868–1873, Townsville, Australia, July 2003.
- M. Villard, M. Trzcinski, and G. Merriam. Fragmentation effects on forest birds: relative influence of woodland cover and configuration on landscape occupancy. *Conservation Biology*, 13: 774–783, 1999.
- B. Williams. sassessment of accuracy in the mapping of vertebrate biodiversity. *Journal of Environmental Management*, 47:269–282, 1996.