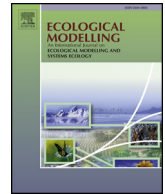




ELSEVIER

Contents lists available at ScienceDirect

## Ecological Modelling

journal homepage: [www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)

# Inferring ecological processes from population signatures: A simulation-based heuristic for the selection of sampling strategies

Benoit Bellot<sup>a,b,\*</sup>, Sylvain Poggi<sup>a</sup>, Jacques Baudry<sup>b</sup>, Yoann Bourhis<sup>c</sup>, Nicolas Parisey<sup>a</sup>

<sup>a</sup> INRA, UMR 1349 IGEPP, Domaine de la Motte, BP 35327, 35653 Le Rheu cedex, France

<sup>b</sup> INRA UMR BAGAP, 65 Rue de St-Brieuc, CS 84215, 35042 Rennes Cedex, France

<sup>c</sup> Rothamsted Research, Department of Biointeraction and Crop Protection, ALS 2JQ Harpenden, UK

## ARTICLE INFO

## Keywords:

Ecological traits  
Reaction–diffusion models  
Population patterns  
Sampling designs  
Classification algorithm  
Landscape ecology

## ABSTRACT

A good knowledge about species traits variability in relation to their environment is the cornerstone of landscape-oriented species management studies. One way to infer this relationship is to compare species signatures in space and time from field data with spatially explicit population dynamics models outputs. However, the inference robustness relies on the available field data, and thus on the quality of the underlying sampling strategy. Field sampling is constrained by several factors, such as the number of landscape replicates, possible number of temporal sessions and number of sample locations, that need to be accounted for prior to field sampling. We set and illustrate a heuristic method to answer the question of optimal sampling conditioned by these landscape-induced constraints. First we studied a real agricultural landscape to determine its mean properties in terms of configuration and composition. The real landscape properties were used as constraints in a landscape model to generate a collection of landscapes with similar properties. On the other hand, we formulated population dynamics models (hereafter noted *Process Models (PM)*) carrying competing hypotheses about two ecological processes—population growth and dispersal—in relation to spatial covariates for *Pterostichus melanarius*, a carabid species involved in pest regulation. We simulated these spatially explicit models and extracted their sampling-dependent signatures, *i.e.* metrics computed on different population samples. We defined a sampling design quality as its ability to capture the contrasts between the *PM* signatures, summarised by the performance of a classification procedure. The most relevant sampling design was selected on the basis of classification performance and *in situ* feasibility. Finally we explored the effects of the *a priori* ecological hypotheses quality on classification performances, through a sensitivity analysis of the *PM* parameters. While some improvements remain to be achieved before being fully operational for landscape ecologists, our framework contributes to bringing closer sampling theory and its application on the field. It endorses the use of landscape modelling to design sampling prior to field experiment to bring out the best from sampled data.

## 1. Introduction

One major ambition of applied ecology is to arrange the landscape structure for a purpose of population dynamics management in space and time (Turner, 2005). This approach serves a wide variety of purposes, such as conservation of endangered species and biodiversity management (Moloney and Levin, 1996; Aben et al., 2016), enhancement of ecosystem services (Landis et al., 2000; Jonsson et al., 2014; Parisey et al., 2016) and control of invasive species (Jonjeans et al., 2008; Savage and Renton, 2014). An increasingly used approach to tackle these thematic is the exploration of the effects of landscape scenarios on the population dynamics through landscape modelling (Schröder and Seppelt, 2006). In this approach, a spatially explicit

population model is simulated on landscapes presenting a gradient of one or several properties to be explored (*e.g.* configurational or compositional heterogeneity, *sensu* Fahrig et al. (2011)). Simulated population data are then characterised by spatial and/or temporal metrics and conclusions are made about the landscape properties that better answer the ecological question at stake (*e.g.* Fahrig, 1998).

For a given species, traits values—*e.g.* movement, reproduction and mortality rates—depend on environmental variables—such as amount and quality of feeding resources, (micro-)climatic conditions—that are spatially heterogeneous at the landscape scale (Honek and Jarosik, 2000; Holland and Luff, 2000). Traits sensitivity to environmental variables however vary from one species to another, and from one trait to another (Jackson and Fahrig, 2015). Spatial heterogeneity of the

\* Corresponding author at: INRA, UMR 1349 IGEPP, Domaine de la Motte, BP 35327, 35653 Le Rheu cedex, France.

E-mail address: [benoit.bellot@gmail.com](mailto:benoit.bellot@gmail.com) (B. Bellot).

<https://doi.org/10.1016/j.ecolmodel.2018.06.010>

Received 25 October 2017; Received in revised form 7 June 2018; Accepted 8 June 2018

0304-3800/© 2018 Published by Elsevier B.V.

environmental variables will thus generate different population patterns according to their impacts on the species life-traits (Fahrig et al., 2015). When attempting to determine the effect of landscape structure through spatially explicit modelling approach, one has to include these processes both qualitatively (*i.e.* which traits have to be included and are they sensitive to environmental heterogeneity?) and quantitatively (what are the traits values with regard to the environmental variables?) (Schröder and Seppelt, 2006). Unfortunately this knowledge is rarely available in the literature, and hardly acquirable through experimentation, both because quantifying movement or reproduction is technically difficult, and that this quantification should be done along the whole environmental gradient. This calls for alternative methods to determine the traits–environment relationships.

While the traits–environmental variables relationships are responsible for the population patterns in space and time, it is reciprocally possible, under some conditions, to infer these processes from the analysis of population patterns (Fortin and Dale, 2005; Grimm et al., 2005). This assumption is a central theme of movement ecology (Avgar et al., 2013; Nathan, 2008) and some studies successfully inferred the movement processes underlying population patterns. McIntire et al. (2013) for instance were able to determine the movement behaviour of a butterfly species by comparing population data from the field and data simulated by individual-based models carrying different mechanistic hypotheses.

Following the concept of “strong inference” initially formulated by Platt (1964), two conditions need to be fulfilled to ensure the robustness of the ecological processes inference: (i) multiple competing hypotheses about the processes must be formulated *a priori* (Johnson and Omland, 2004), and (ii) the experimental design effectively allows discriminating the competing hypotheses (McIntire and Fajardo, 2009; Platt, 1964). In the discipline of landscape ecology, where the experiment consists in sampling individuals in space and time, this second condition means that the sampling must be designed such as the population patterns associated to different processes are contrasted enough to be distinguishable (Schweiger et al., 2016; Schröder and Seppelt, 2006). Using such a design on the field thus ensures that the extracted pattern leads to the correct selection of the underlying ecological process.

Conversely, insufficient sampling may lead to a pattern description that is too coarse to disentangle the competing hypotheses formulated about the ecological processes (McIntire and Fajardo, 2009). The question of optimal sampling design to unravel ecological processes from the population pattern has thus been of an increasing interest for the last four decades (Mateu and Müller, 2012). In this context, Zurell et al. (2010) emphasised the exciting potential of the use of virtual population data to evaluate methods for data sampling, naming this approach as the “virtual ecologist” methodology. However, to the best of our knowledge, no methodological framework using simulated population data has been proposed to set up sampling over space and time with the specific objective of inferring the ecological processes driving a species dynamics from a set of *a priori* hypotheses. Defining the population signature as the censored information resulting from the sampling of a population spatio-temporal patterns, and based on the idea that ecological processes can be inferred from this signature (Tilman and Kareiva, 1997; Tuda, 2007), we propose a heuristic framework to address this methodological issue.

After presenting the general approach of our framework, we illustrate its application through a real landscape based scenario to define *a priori* how sampling should be conducted to infer the ecological processes driving the population dynamics of a ground-dwelling carabid species, *Pterostichus melanarius*, among a set of competing hypotheses.

Although *P. melanarius* population distributions within the agroecosystem have been widely studied (*e.g.* Holland and Begbie, 2004; George Thomas and Holland, 2002), considerable gaps in our knowledge remain and must be filled if we are to manipulate the farm habitat in its favour (Holland and Luff, 2000). For example, knowledge related

to its preferred habitat types remains unclear, as well as the processes governing its heterogeneous distribution within and between fields (Holland and Luff, 2000). Predicting population patterns in relation to landscape structure through an approach of landscape modelling cannot be conducted without such ecological knowledge (Jongejans et al., 2008; Schröder and Seppelt, 2006).

We tackle this lack of knowledge about this key natural enemy species, by formulating competing mechanistic hypotheses that are susceptible to drive its population dynamics within the agricultural landscape. The goal of our study was to identify *a priori* sampling designs that would best allow selecting the correct ecological process at stake from the sampled data. In turn, we did not infer the ecological processes driving *P. melanarius* population dynamics at the landscape scale. Rather we proposed a sampling design to do so in a real agricultural study area, within the range of sampling efforts usually made for studies where carabid are sampled within agricultural landscapes. The outputs of our study are (i) a proposition of sampling design in terms of number of sampling points, number of sampling sessions in time, and number of landscape replicates in which to sample, (ii) the associated classification model to be used for ecological process inference on the basis of sampled data and (iii) the assessment of the population parameters uncertainty on the inference robustness, through a sensitivity analysis framework. Finally we discuss the implications of such framework, and what improvements remain to be done before being fully operational for ecologists that aim to infer ecological processes from the population signatures at the landscape scale.

## 2. Material and methods

### 2.1. General heuristic approach

The general approach of our heuristic is illustrated in Fig. 1. It is designed to address the question of the optimal sampling strategies in order to maximise the ability to determine the biological-traits/environment relationships of the species of interest. The approach involves four steps. The first step consists in generating virtual landscapes presenting realistic structural properties. The study area that is chosen for biological sampling is analysed and mean structural properties of the landscapes are extracted. These properties are included as parameter values in a landscape model to generate spatial domains with targeted properties. The second step is devoted to the generation of virtual population data. It consists of formulating competing ecological hypotheses from the available knowledge about the species of interest, translating them into mechanistic, spatially explicit population dynamics models (*Process Models, PMs*), and simulating them on the virtual landscapes. In the third step, sampling designs differing in their spatial and temporal components  $S_s, S_t, S_r$  (respectively the numbers of sampling locations, sampling dates and landscape replicates) are applied to sample the simulated population data. A set of summary statistics (hereafter named *signature*  $\Phi$ ) are computed on sampled data for each sampling design, defining the sampling dependent *PM* signatures. For each sampling design, the couples (*PM*,  $\Phi$ ) are used as inputs for a classification model that learns and predicts the *PM* as a function of their population signatures. The classification accuracy, defined as the rate of correct predictions of the *PMs* from their signatures, quantifies the sampling design quality: a high accuracy means that the sampling design generates *PM*-specific signatures that are easily distinguishable one from another. Using a classification model also allows to extract other information about the classification performances through the analysis of the confusion matrices for each sampling design, such as type I and II error rates (respectively “false positives” and “false negatives”) for each *PM*, or their pairwise confusion, which is a proxy towards *PM* signatures similarity. Such information quantifies the uncertainty associated with the classification of ecological processes from sampled data. Another valuable information is to describe the importance of the sampling components ( $S_s, S_t, S_r$ ) on classification

Download English Version:

<https://daneshyari.com/en/article/8845948>

Download Persian Version:

<https://daneshyari.com/article/8845948>

[Daneshyari.com](https://daneshyari.com)