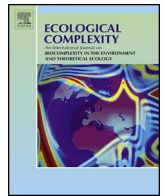




Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom



Original Research Article

‘Catch me if you can’: Evaluating the population size in the presence of a spatial pattern

N.B. Petrovskaya

School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:

Received 10 December 2016
Received in revised form 9 March 2017
Accepted 17 March 2017
Available online xxx

Keywords:

Population size
Population density
Spatial density distribution
Sampling plan
Sampling grid
Coarse grid
Sparse data

ABSTRACT

Many biological and ecological problems require accurate evaluation of the total population size. We discuss a sampling procedure used for evaluation of the population abundance from information collected on a grid of spatial sampling locations. It is shown in the paper how insufficient information about the spatial population density obtained on a coarse sampling grid affects the accuracy of evaluation. The insufficient information is collected because of inadequate spatial resolution of the population density on coarse grids and this is especially true when a heterogeneous spatial population is sampled. It is argued in the paper that the evaluation error is a random variable on coarse sampling grids because of the uncertainty in sampling spatial data and a probabilistic approach should be employed in the evaluation procedure. We also show that there exists a threshold number of sampling locations on a regular sampling grid where we can guarantee desired accuracy of evaluation. Information about the threshold number of sampling locations allows one to reconcile the probabilistic approach based on the assumption about randomness of sampling data with the deterministic approach based on the requirement that spatial data are collected only once as the sampling procedure cannot be repeated under the same conditions.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Many biological and ecological problems require evaluation of the total population size. This evaluation should be accurate as insufficient information about the total population size may make an undesirable impact on the ecosystem, e.g. unnecessary application of pesticide if the pest abundance is not correctly evaluated (Jepson and Thacker, 1990). Inaccurate evaluation of the total population size may also result in a wrong conclusion about presence or absence of some important ecological traits, e.g. synchronization between population fluctuations in different habitats (Petrovskaya and Petrovskii, 2017).

A standard approach to evaluation of the total population size is to consider a simple estimate of the spatial population density integral (Davis, 1994; Snedecor and Cochran, 1980) where a sampling grid is used to collect data related to the spatial density distribution. The definition of a sampling grid depends on an ecological problem where sampling is required. For an estimate of the total population size to be accurate the sampling grid must capture sufficient information to adequately represent the true

population size. One important consideration is the sampling plan, that is, the prescribed locations at which samples are to be taken. Comparisons of various spatial arrangements have been made in order to make recommendations (Alexander et al., 2005) but in many cases sampling locations are defined as nodes of a regular grid (Ferguson et al., 2000; Holland et al., 1999).

Another important factor is the total number of locations where samples are taken. In some cases this number is derived based on theoretical recommendations (Taylor et al., 1978), while in many other cases this property of the sampling grid is decided *ad hoc* (Boag et al., 2010). For example, a widely used sampling technique is trapping where trap counts provide information about the population density at the position of the traps (Byers et al., 1989; Raworth and Choi, 2001). In a trapping procedure applied in routine insect pest monitoring programs the number N of traps rarely exceeds twenty per a typical agricultural field with a linear size of several hundred meters (Mayor and Davies, 1976) and in some cases N can be as small as one or a few traps per field (Northing, 2009).

Sampling protocols do not usually allow one to make extensive repetition of sampling. Sometimes a pre-sample (or series of them) can be used to obtain a sample mean and sample variance from which an estimate of the number of sample units needed to

E-mail address: n.b.petrovskaya@bham.ac.uk (N.B. Petrovskaya).

achieve a specified accuracy can be calculated; e.g. see Binns et al., 2000; Dent, 2000; Pedigo and Rice, 2009. However there always is a trade-off between the number of sample units needed to achieve sufficient accuracy and the number that can be practically afforded. Moreover even when the practitioners are not happy with the accuracy of the sampling procedure a repeated sampling with an increased number of samples is not available in most ecological applications due to the impossibility to reproduce the same sampling conditions in the environment (the air temperature, the soil moisture, etc.).

The above restrictions on a sampling protocol lead to a question about accuracy of evaluation of the total population size on coarse sampling grids. It has been shown in our previous work (Embleton and Petrovskaya, 2013; Petrovskaya and Embleton, 2013, 2014) that the standard evaluation technique does not work when coarse grids are used for evaluation because of the insufficient information (uncertainty) in data collected on such grids. It was suggested in Petrovskaya and Embleton (2013) that the total population size on coarse sampling grids is a random variable and it has to be handled by using probabilistic techniques. However, while a probabilistic approach implies that multiple realisations of a random variable are available in the evaluation task, the sampling procedure is 'deterministic' because it has to deal with a single realisation of the total population size based on data collected only once. How can then practitioners know that a single realisation of the random variable gives them a sufficiently reliable result? In the present paper we discuss this issue and introduce reliability criterion required to 'reconcile' the probabilistic and deterministic approaches when the total population size is evaluated.

The paper is organized as follows. In Section 2 we briefly revisit the problem of evaluation of the total population size and introduce the concept of uncertainty in sampling data related to a spatial pattern of the density distribution. We discuss how insufficient information about the spatial pattern can be re-formulated in terms of location of grid nodes on a coarse sampling grid. We then consider the evaluation error as a random variable and calculate probabilistic characteristics of the total population size in Section 3. We discuss the probability of accurate evaluation of the population size on a sampling grid with a given number of grid nodes in Section 4. It will be argued in Section 4 that probabilistic characteristics of the evaluation error become insignificant when a sampling grid is fine enough to guarantee that the desired accuracy can be achieved with the probability equal to one. Thus transition from a probabilistic problem to a deterministic problem can be formulated in terms of a certain threshold number of grid nodes on a sampling grid. Several numerical examples are provided to illustrate the above argument. Finally, in Section 5 we summarise our experience with the problem of accurate evaluation of the total population size.

2. Evaluation of the total population size on coarse sampling grids

In the present paper we do not refer to any particular sampling technique yet we assume that samples are collected at nodes of a regular spatial grid. This assumption can be readily linked to sampling procedures used in real-life ecological applications such as e.g. pitfall trapping of invertebrates (Woodcock, 2008). Furthermore we assume that sampling brings us reliable information about the population density $u(x,y)$, i.e. we have exact values $u(x_i, y_i)$ of the population density at sampling locations (x_i, y_i) , $i = 1, \dots, N$. The further discussion of the latter assumption can be found in Bearup et al. (2015).

If the population density function $u(x, y)$ were known at any point (x, y) of the domain D then the total population size I would be given by

$$I = \iint_D u(x, y) dx dy. \quad (1)$$

However we have a discrete density distribution as a result of sampling and we therefore have to replace integral (1) with a weighted sum of density values. The following formula is widely used in practical ecological applications, e.g. see Davis (1994) and Snedecor and Cochran (1980),

$$I \approx I_a(N) = \frac{A}{N} \sum_{i=1}^N u_i, \quad (2)$$

where I_a is the approximate value of the total population size I and A is the area of the domain. It is important to note that the result of evaluation I_a depends explicitly on the number N of sampling locations, $I_a = I_a(N)$.

The evaluation error has to be introduced since the exact population size is replaced by some approximation I_a . The relative evaluation error is defined as

$$e(N) = \frac{|I - I_a(N)|}{I}, \quad (3)$$

where we assume that the exact population size is $I > 0$. Conclusions about accuracy of evaluation can then be made based on the following requirement

$$e(N) \leq \tau, \quad (4)$$

where τ is specified tolerance. In ecological applications the accuracy requirement (4) is not very demanding as the tolerance $0.2 < \tau < 0.5$ is already considered as acceptable (Pascual and Kareiva, 1996; Sherratt and Smith, 2008). However we will see later in the paper that even this relatively high tolerance cannot always be provided.

One important feature of approximation (2) is convergence of I_a to the exact value I when the hypothetical number of sampling locations is very large. We have the absolute error and the relative error of evaluation

$$|I - I_a(N)| \rightarrow 0 \text{ as } N \rightarrow \infty, \quad e(N) \rightarrow 0 \text{ as } N \rightarrow \infty. \quad (5)$$

At the first glance, the choice of method (2) may not be optimal for evaluation of total population size (1). From a numerical integration viewpoint the method (2) can be loosely interpreted as the midpoint rule of integration where the whole domain is subdivided into N subdomains with the area of each subdomain given by A/N and the density function $u(x)$ is approximated by constant u_i , $i = 1, 2, \dots, N$ in each of those subdomains. Convergence properties (5) of the midpoint rule are well studied and it also is well known, e.g. see Davis and Rabinowitz (1975) that its convergence rate is inferior to more advanced methods of numerical integration, e.g. the Simpson method on regular grids or Gauss quadratures Petrovskaya et al. (2012). However, the aim of this paper is to study the accuracy of the total population size evaluation on coarse sampling grids where the number N of sampling locations is small. It will be shown later in the paper that accuracy of evaluation on coarse grids is not related to the convergence rate of the method when N is large. Hence methods with higher convergence rate cannot be immediately recommended for the purpose of our study and for the sake of clarity we choose to consider a simple procedure (2) instead of an advanced method of numerical integration in order to introduce our approach to accuracy evaluation on coarse grids.

Coarse sampling grids are typical in many sampling protocols because financial and labor resources available for sampling are

Download English Version:

<https://daneshyari.com/en/article/8844826>

Download Persian Version:

<https://daneshyari.com/article/8844826>

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