Ecological Complexity 21 (2015) 185–198

Contents lists available at ScienceDirect

Ecological Complexity

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

On the accuracy of estimating pest insect abundance from data with random error

Nina Embleton *, Natalia Petrovskaya

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

ARTICLE INFO

Article history:
Received 3 January 2014
Received in revised form 5 May 2014
Accepted 11 May 2014
Available online 4 July 2014

Keywords:
Pest insect monitoring
Noise
Numerical integration
Trapezoidal rule
Simpson's rule

ABSTRACT

Numerical integration is a popular technique that can be successfully applied to evaluating the pest insect abundance in an agricultural field. In this paper we apply numerical integration in the problem where data about insects obtained as a result of a trapping procedure have random error (noise). We compare several methods of numerical integration that have different accuracy of evaluation when precise data are considered. In particular, we consider the composite trapezoidal and composite Simpson's rules of integration, and compare them with a statistical approach to obtaining an estimate based on the sample mean. The comparison is first done in the case when the number of traps where the data are available is large. It will be shown in the paper that noise in the data badly affects the accuracy of evaluation on fine grids of traps, so the different methods of numerical integration no longer differ in terms of their accuracy. We then consider an ecologically relevant case of a small number of traps, *i.e.* when the data available for evaluation are sparse. It will be discussed in the paper that the impact of noise is negligible on coarse grids of traps and therefore we can keep the accuracy hierarchy of numerical integration methods established from the consideration of precise data. We are then able to give recommendations on how to use methods of numerical integration to evaluate pest abundance. Our results are illustrated by numerical experiments.

Nomenclature

$\tilde{E}_{\max}$	upper limit of the credible interval of $\tilde{E}_{rel}$
$\tilde{E}_{\min}$	lower limit of the credible interval of $\tilde{E}_{rel}$
E_{rel}	relative error of the estimate I_a (noise is absent)
$\tilde{E}_{rel}$	relative error of the estimate $\tilde{I}$ (noise is present)
f	pest population density function
I	exact pest abundance
$\tilde{I}$	estimate of pest abundance formulated from noisy density data
I_a	estimate of pest abundance formulated from exact density data
$u(\tilde{I})$	uncertainty associated with the estimate $\tilde{I}$
$\mu(\tilde{E}_{rel})$	mean of the error quantity $\tilde{E}_{rel}$
$\sigma_{\tilde{I}}$	standard deviation of the estimate $\tilde{I}$

1. Introduction

Accurate evaluation of pest insect abundance is a key component in any integrated pest management (IPM) programme used in agriculture (Burn et al., 1987; Metcalf and Luckmann, 1982). The decision of whether or not to implement a control action to manage the pest population is made by comparing an estimate to some threshold value(s) (Stern, 1973; Stern et al., 1959). The decision can be considered to be correct if the same conclusion would have been reached if the true abundance had been known. However, by definition of the problem the true abundance is unknown, thus we require information about the reliability of the estimate in order to have confidence about the management decision. Knowledge of the accuracy of an estimate can give us an indication of the relationship between the true pest abundance and the threshold value(s) and thus we can establish if there is a risk of an incorrect decision. The risk grows smaller as the estimate becomes more accurate.

Evaluation is based on the results of sampling and its accuracy depends on a sampling technique. Trapping is a sampling procedure widely employed in monitoring. The idea is that trap counts can be converted into the pest population density at trap locations in order to obtain an estimate of the total pest population

* Corresponding author.
E-mail addresses: embleton@for.mat.bham.ac.uk (N. Embleton), n.b.petrovskaya@bham.ac.uk (N. Petrovskaya).

http://dx.doi.org/10.1016/j.ecocom.2014.05.006
1476-945X/

size (Byers et al., 1989; Raworth and Choi, 2001). The accuracy of such evaluation depends strongly on how the data are collected and the crucial factor with regard to data collection is the number of traps available in the monitoring procedure. Under routine monitoring, financial conditions and other restrictions do not normally allow for a big number of traps and that, in turn, may result in poor accuracy of evaluation.

Apart from the methodology of data collection another important issue is how the trap counts are processed. Methods of numerical integration are a well-known family of methods designed to handle discrete data (Davis and Rabinowitz, 1975). Their application in the pest insect monitoring problem has been studied in Embleton and Petrovskaya (2013), Petrovskaya and Embleton (2013), Petrovskaya and Petrovskii (2010), Petrovskaya et al. (2012, 2013), and Petrovskaya and Venturino (2011). It was discussed in Petrovskaya and Embleton (2014) that the application of more advanced numerical integration techniques often results in a more accurate evaluation of pest abundance than straightforward statistical computation of the mean density, *cf.* Davis (1994) and Snedecor and Cochran (1980).

The initial study of numerical integration techniques for the pest abundance evaluation problem has been made under the assumption that density data obtained as a result of trapping are precise. The above assumption is not entirely realistic and the results should therefore be extended to the case when the density measurements have random error. The measurements of density are thus associated with some *uncertainty* rather than being definitively known quantities and this gives rise to uncertainty in the abundance estimate and in turn in the accuracy of this estimate. It is important to mention that the measurements obtained *via* trapping are also dependent on the activity of the target species as well as their density. In order to truly reflect the density, the measurements must be calibrated somehow (Petrovskii et al., 2012; Raworth and Choi, 2001). This calibration induces another error into the estimate, however, within this paper we ignore this error. Instead, we assume that the measurements already reflect the pest density but that there is some additional random error (noise) present.

The accuracy of a selected method of numerical integration (the trapezoidal rule) applied to data measured with random error has been investigated in our recent paper (Embleton and Petrovskaya, 2014). It was shown in Embleton and Petrovskaya (2014) that the results of numerical integration of noisy data depends strongly on the number of traps where the data are collected. Namely, if the number of traps is large, noise becomes a dominant feature of the pest abundance approximation and the results may differ from an estimate of the pest abundance obtained from precise data by several orders of magnitude. On the other hand, noise does not have a lot of impact on the accuracy of a pest abundance estimate when the number of traps is small.

As we have already mentioned, the conclusions of the paper (Embleton and Petrovskaya, 2014) concern the trapezoidal rule of integration only. Meanwhile, it is possible to employ a different method of numerical integration to evaluate the total pest population size. The results of Petrovskaya and Embleton (2014) and Petrovskaya et al. (2012) have revealed that so-called higher order methods of integration provide better accuracy when exact data are considered. Thus the question arises if higher order methods will have an advantage in accuracy when the pest abundance is approximated based on noisy data and this question is the focus of the present paper.

Keeping in mind the results of our previous study (Petrovskaya and Embleton, 2014; Petrovskaya and Petrovskii, 2010), the question of accuracy must be investigated separately for the case of a small number of traps (*i.e. coarse grids* of traps) and a large number of traps (*fine grids*), as different approaches have to be applied in order to validate the accuracy in the former and latter case. Hence the paper is organised as follows. In the next section, we briefly explain basic numerical integration techniques under the assumption that an estimate of pest abundance is based on precise data. In Section 3 we recall the results of our paper (Embleton and Petrovskaya, 2014) to establish how random error in data translates to error in a pest abundance estimate. We then apply the results of Section 3 to compare three methods of numerical integration on fine grids in Section 4, where the convergence rate of the mean error is discussed. The same methods of numerical integration are compared on coarse grids in Section 5. The results of previous sections are illustrated by designed numerical examples in Section 6 for ecologically relevant test cases. Finally, concluding remarks are provided in Section 7.

2. Numerical integration as a means of estimating pest abundance

In this section we discuss the implementation of numerical integration methods within the framework of pest monitoring. For the sake of simplicity, we reduce the problem to one dimension and essentially consider an agricultural field as a straight line. Let us note, however, that the results of our study can readily be expanded to multi-dimensional problems.

Once information on the pest population in an agricultural field has been collected by some chosen means of sampling, an estimate of the abundance can be formed. Typically the estimate used within the ecological community depends on the sample mean (Davis, 1994). Counts obtained from sampling can be manipulated to give the pest density at each sample unit location (Byers et al., 1989; Raworth and Choi, 2001). We shall use the notation f_i to denote the pest population density at the sample unit location x_i, $i = 1, \ldots, N$. An estimate I_a to the true abundance I can be calculated thusly

$$I \approx I_a = L\bar{f}, \quad \bar{f} = \frac{1}{N}\sum_{i=1}^{N} f_i,$$

where L is the length of the field, $\bar{f}$ is the sample mean pest density, and N is the total number of sample units. Let the domain of the agricultural field be further represented by the unit interval [0, 1], since a simple linear transformation can be applied to yield an interval of arbitrary length L. The above estimate of the abundance then becomes equivalent to the sample mean pest density, namely,

$$I \approx I_a = \frac{1}{N}\sum_{i=1}^{N} f_i. \quad (1)$$

The formula (1) calculates an estimate of the pest insect abundance as a weighted sum of the density function values. This approach can be further generalised to arrive at a family of numerical integration methods as discussed in Petrovskaya and Embleton (2014). Theoretically speaking, the exact pest population abundance I could be obtained by integrating analytically the pest population density function $f(x)$,

$$I = \int_0^1 f(x)dx,$$

if we knew a continuous density function $f(x)$ on the interval [0, 1]. In reality, however, information on the pest density is provided by sampling the population and the population density function is consequently discrete, namely, $f(x) \equiv f_i$, $i = 1, \ldots, N$. The above integral thus cannot be evaluated and we must instead seek an approximation I_a to the exact pest abundance I by means of numerical integration.