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

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,

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

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.

The general formula for numerical integration is given by the weighted sum (*e.g.* see (Davis and Rabinowitz, 1975))

$$I \approx I_a = \sum_{i=1}^{N} w_i f_i \tag{2}$$

where the weights $w_i, i = 1, \ldots, N$ depend on the specific method of numerical integration chosen to be employed. It is easy to see on comparing the formula (2) with the definition of the sample mean density (1) that this estimate can be considered as a simple form of a numerical integration method where the weights are uniformly defined as

$$w_i = \frac{1}{N}, \quad i = 1, \ldots, N. \tag{3}$$

There are of course many other possible combinations of weight coefficients which can be used in the formula (2) to yield an estimate $I_a$. The scope of this paper is restricted to the consideration of just three particular examples of weight coefficients as explained below.

Alongside the sample mean density (3) we study estimates formed from two members of the composite Newton–Cotes family of numerical integration methods (*e.g.* see Davis and Rabinowitz (1975)). To apply a method belonging to this family the locations $x_i$ at which the function values $f_i$ are available are required to be regularly spaced. Such a sampling plan is indeed often used in pest monitoring (Ferguson et al., 2000; Holland et al., 1999). The weights of a composite Newton–Cotes formula are derived by piecewise polynomial interpolation of the discrete function $f(x) \equiv f_i$, $i = 1, \ldots N$ for a chosen degree of interpolating polynomial. The composite trapezoidal rule is formed by piecewise linear polynomial interpolation and has the weights

$$w_i = h, \quad i = 2, \ldots, N-1, \qquad w_i = \frac{h}{2}, \quad i = 1, \text{ or } i = N. \tag{4}$$

where $h = 1/(N-1)$ is the fixed distance between sampling points. The number $N$ of traps is required to satisfy the condition $N \geq 2$.

Another method we discuss in the paper is the composite Simpson's rule. This integration technique uses piecewise quadratic polynomial interpolation and requires an additional restriction to be imposed, namely, that the number of sampling locations $N$ must be odd and such that $N \geq 3$. The weights are described by

$$\begin{aligned} w_i &= \frac{4h}{3}, \quad i = 2, 4, \ldots, N-1, \\ w_i &= \frac{2h}{3}, \quad i = 3, 5, \ldots, N-2, \qquad w_i = \frac{h}{3}, \quad i = 1, \text{ or } i = N. \end{aligned} \tag{5}$$

In order to make comparisons between the methods (3)–(5) we henceforth consider the following regular distribution of the sampling locations $x_i$ across the unit interval:

$$x_i = \frac{i-1}{N-1}, \quad i = 1, \ldots, N, \tag{6}$$

where the number $N \geq 3$ of sampling points is odd.

If the exact pest abundance $I$ is known then the accuracy of an estimate $I_a$ can be assessed by considering the *approximation error*. Since the pest abundance is $I > 0$, the relative approximation error $E_{rel}$ is defined as

$$E_{rel} = \frac{|I - I_a|}{I}. \tag{7}$$

Clearly the smaller the relative error, the more accurate the corresponding estimate $I_a$. A more accurate estimate gives rise to greater confidence that the correct decision of whether or not to implement a control action can be made. Therefore, we impose the following condition on the relative error $E_{rel}$:

$$E_{rel} \leq \tau, \tag{8}$$

for some specified accuracy tolerance $\tau$. In ecological applications a tolerance $\tau$ such that $\tau \in [0.2, 0.5]$ is considered acceptable (Pascual and Kareiva, 1996; Pedigo and Rice, 2009).

## 3. The uncertainty introduced by random error

A trap count can be manipulated to provide a measurement of the pest density at the trap location, however, a measurement is subject to *measurement error*. Let us denote the measured pest density at the trap location $x_i$ by $\tilde{f}_i$, and now $f_i$ is used to represent the corresponding exact pest density. The relationship between the measured pest density $\tilde{f}_i$ and the true pest density $f_i$ is then $f_i = \tilde{f}_i + \epsilon_{m_i}$ where $\epsilon_{m_i}$ is the measurement error. A measurement error is considered to consist of two components: a random component, and a systematic component (BIPM, 2008). In other words, the measurement error $\epsilon_{m_i}$ can be expressed as $\epsilon_{m_i} = \epsilon_{r_i} + \epsilon_{s_i}$ where $\epsilon_{r_i}$ and $\epsilon_{s_i}$ represent the random and systematic error, respectively. The random error is the result of noise in the data and thus any $\epsilon_{r_i}, i = 1, \ldots N$ can be either positive or negative with equal probability. The systematic error on the other hand is caused by some source of bias and therefore every $\epsilon_{s_i}, i = 1, \ldots N$ is consistently either positive or negative. This paper focuses on the impact of noise in the data and as such we ignore the systematic contribution to the measurement error. That is to say we redefine the relationship between the measured quantity $\tilde{f}_i$ and the true value $f_i$ as

$$f_i = \tilde{f}_i + \epsilon_{r_i}.$$

The random error component $\epsilon_{r_i}$ of a measured pest density $\tilde{f}_i$ given in the above equation is in essence a realisation of a random variable. We consider the true pest density $f_i$ to be some unknown constant. Since $\tilde{f}_i$ is the sum of an unknown constant $f_i$ and a realisation $\epsilon_{r_i}$ of a random variable, it can in turn also be considered a realisation of another random variable. There is thus an *uncertainty* associated with a measured pest density $\tilde{f}_i$. We follow the procedure outlined in our previous work (Embleton and Petrovskaya, 2014) and consider each $\epsilon_{r_i}$ to be a realisation of a normal distribution where we make the assumption that the mean is zero. This means each $\tilde{f}_i, i = 1, \ldots N$ is thus a realisation of a normally distributed random variable $F_i$ with mean $\mu_i = f_i$ and standard deviation $\sigma_i$. Furthermore, since each $\tilde{f}_i$ is a realisation of a normally distributed random variable $F_i$, a measurement $\tilde{f}_i$ belongs to the range

$$\tilde{f}_i \in [\tilde{f}_i^{\min}, \tilde{f}_i^{\max}] = [f_i - z\sigma_i, f_i + z\sigma_i] \tag{9}$$

with probability

$$P(z) = erf\left(\frac{z}{\sqrt{2}}\right), \tag{10}$$

where $erf(z)$ is the error function.

To define the standard deviation $\sigma_i$ we further impose the condition that the range of each of the measured quantities $\tilde{f}_i$ shall be restricted as belonging to the following interval with probability $P(z)$:

$$\tilde{f}_i \in [\tilde{f}_i^{\min}, \tilde{f}_i^{\max}] = [(1-\nu) f_i, (1+\nu) f_i], \tag{11}$$

where $\nu \in (0, 1)$ is the measurement tolerance defined by the conditions of the experiment. Equating the ranges (9) and (11) gives an expression for the standard deviation $\sigma_i$ as

$$\sigma_i = \frac{\nu f_i}{z}. \tag{12}$$

We note that the definition $\tilde{f}_i$ does not depend on how long traps are exposed in the field (see our previous work (Embleton and Petrovskaya, 2014) for further discussion of this topic).

Applying a method of numerical integration (2) to a measured data set $\tilde{f}_i, i = 1, \ldots, N$ instead of the exact values $f_i$, $i = 1, \ldots, N$ yields an estimate based on measured data $\tilde{I}$:

$$\tilde{I} = \sum_{i=1}^{N} w_i \tilde{f}_i. \tag{13}$$

We recall that $I > 0$, thus the relative approximation error of such an estimate which we denote $\tilde{E}_{rel}$ is defined as

$$\tilde{E}_{rel} = \frac{|I - \tilde{I}|}{I}. \tag{14}$$

It can readily be seen from (13) and (14) that both the estimate $\tilde{I}$ and the corresponding relative error $\tilde{E}_{rel}$ depend on the measured values $\tilde{f}_i, i = 1, \ldots, N$. It follows that the uncertainty associated with these measured pest densities will give rise to uncertainty in the quantities $\tilde{I}$ and $\tilde{E}_{rel}$. In our previous work (Embleton and Petrovskaya, 2014) we established the credible interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ to which the relative approximation error $\tilde{E}_{rel}$ of an estimate $\tilde{I}$ belongs with probability $P(z)$. Below we briefly summarise these results as they are important for our further discussion. Meanwhile, the interested reader is referred to the text (Embleton and Petrovskaya, 2014) for the detailed formulation and justification.

The standard deviation of a random variable provides a measure of its associated uncertainty. We recall that we consider a measured pest population density $\tilde{f}_i$ to be a realisation of a normally distributed random variable $F_i$ with mean $\mu_i = f_i$ and standard deviation $\sigma_i$, where $\sigma_i$ is given by (12). The uncertainty associated with a measured value $\tilde{f}_i$ which we will denote $u(\tilde{f}_i)$ is quantified thusly:

$$u(\tilde{f}_i) = \sigma_i.$$

From (13) it can be seen that the estimate of pest abundance $\tilde{I}$ is a linear combination of the measured densities $\tilde{f}_i, i = 1, \ldots, N$. As such, an estimate $\tilde{I}$ is also a realisation of a normally distributed random variable $\tilde{I}_F$, where

$$\tilde{I}_F = \sum_{i=1}^{N} w_i F_i.$$

The random variable $\tilde{I}_F$ has mean $\mu_{\tilde{I}} = I_a$ and standard deviation $\sigma_{\tilde{I}}$. As explained in Cox (2007) the uncertainty $u(\tilde{I})$ associated with the estimate $\tilde{I}$ can be obtained by propagating the uncertainty associated with the measured values $u(\tilde{f}_i), i = 1, \ldots, N$ by means of the *law of the propagation of uncertainty* (BIPM, 2008). Assuming the data are uncorrelated we arrive at

$$u(\tilde{I}) = \sqrt{\sum_{i=1}^{N} w_i^2 u^2(\tilde{f}_i)}, \tag{15}$$

and we have $u(\tilde{I}) = \sigma_{\tilde{I}}$.

Let us now consider the relative error quantity $\tilde{E}_{rel}$ as given by (14). Clearly, this quantity is the absolute value of a linear function of $\tilde{I}$. As detailed above, $\tilde{I}$ is a realisation of a normally distributed random variable, hence, $\tilde{E}_{rel}$ is in turn a realisation of a random variable $\tilde{E}_F$ with a *folded normal distribution* (Leone et al., 1961). Expressions for the mean and variance of such a distribution are known and standard formulas are given in Leone et al. (1961). Using these results, we are able to give the mean of the integration error as

$$\mu(\tilde{E}_{rel}) = \left(1 - \frac{I_a}{I}\right)\left[1 - 2\Phi\left(\frac{I_a - I}{\sigma_{\tilde{I}}}\right)\right] + \frac{\sigma_{\tilde{I}}}{I}\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{1}{2}\left(\frac{I_a - I}{\sigma_{\tilde{I}}}\right)^2\right\}, \tag{16}$$

where $\Phi$ is the standard normal cumulative distribution function. The standard deviation is

$$\sigma_{\tilde{E}_{rel}} = \sqrt{\left(1 - \frac{I_a}{I}\right)^2 + \left(\frac{\sigma_{\tilde{I}}}{I}\right)^2 - \mu^2_{\tilde{E}_{rel}}}.$$

In our previous work (Embleton and Petrovskaya, 2014) we showed that a realisation $\tilde{E}_{rel}$ of the random variable $\tilde{E}_F$ belongs to the range $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ with the probability $P(z)$ as defined by (10). The range $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ is called the credible interval of $\tilde{E}_{rel}$ (*e.g.* see Bolstad (2007)). The lower limit $\tilde{E}_{\min}$ of the credible interval is given by

$$\tilde{E}_{\min} = \begin{cases} 0, & \text{for } E_{rel} \le z\frac{\sigma_{\tilde{I}}}{I}, \\ E_{rel} - \frac{z\sigma_{\tilde{I}}}{I}, & \text{for } E_{rel} > z\frac{\sigma_{\tilde{I}}}{I}. \end{cases} \tag{17}$$

The upper limit $\tilde{E}_{\max}$ is defined as

$$\tilde{E}_{\max} = \begin{cases} E_{rel} + \frac{\sigma_{\tilde{I}}}{I}\Phi^{-1}\left[2\Phi(z) - \Phi\left(z + \frac{2IE_{rel}}{\sigma_{\tilde{I}}}\right)\right], & \text{for } E_{rel} \le z\frac{\sigma_{\tilde{I}}}{I}, \\ E_{rel} + \frac{\sigma_{\tilde{I}}}{I}\Phi^{-1}\left[\begin{array}{c}\Phi(z) - \Phi\left(z - \frac{2IE_{rel}}{\sigma_{\tilde{I}}}\right) \\ -\Phi\left(z + \frac{2IE_{rel}}{\sigma_{\tilde{I}}}\right) + 1\end{array}\right], & \text{for } E_{rel} > z\frac{\sigma_{\tilde{I}}}{I}, \end{cases} \tag{18}$$

where $\Phi$ is the standard normal cumulative distribution function and $\Phi^{-1}$ is its inverse.

## 4. The accuracy on integration of noisy data: fine grids

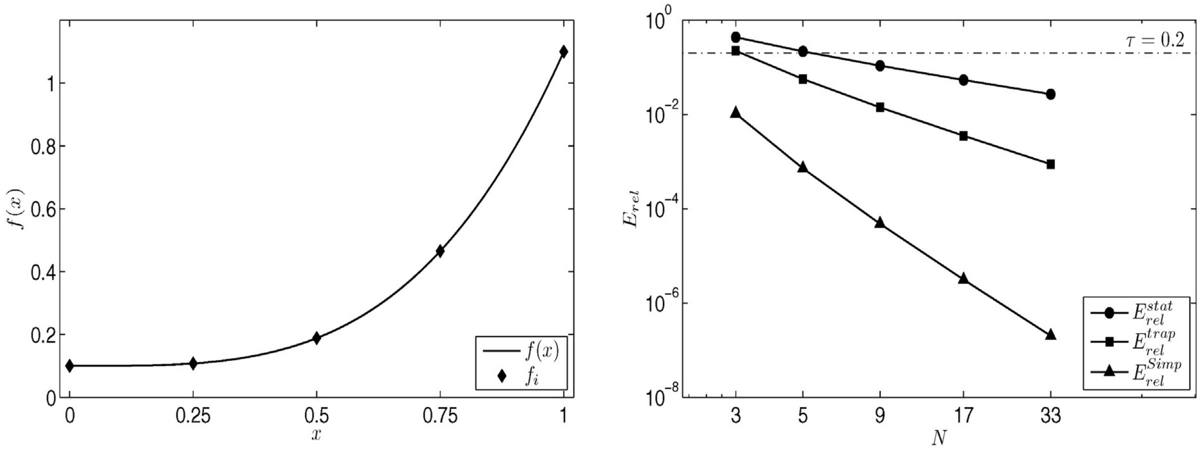
In this section we analyse the formula (16) along with the credible interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ of $\tilde{E}_{rel}$ for the case when the number $N$ of traps where data are available is large. It can be seen from (17) and (18) that the lower and upper bounds of the interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ induced by noise in the pest population density data depend on the accuracy $E_{rel}$ of evaluation obtained when exact density values are considered. Thus we first have to discuss a pest abundance estimate calculated from precise data and below we recall the concept of convergence for the relative error $E_{rel}$.

### 4.1. Convergence on fine grids

Consider an estimate $I_a$ and the corresponding relative error $E_{rel}$ obtained as a result of pest abundance approximation using precise data. It follows from (2) that the error $E_{rel}$ depends on the number $N$ of sample units taken. In order to be a viable method of numerical integration as $N$ increases, the error must decrease. That is to say we have convergence of the relative error to zero, as we increase the number of traps,

$$E_{rel} \to 0 \text{ as } N \to \infty.$$

The rate at which this convergence occurs depends on the choices of the weights $w_i$ in the formula (2). Exploitation of the convergence rate could prove a useful tool in the evaluation of pest abundance. As explained in Section 1, the nature of the pest



(a) (b)

**Fig. 1.** Evaluating the pest abundance by means of numerical integration. (a) A toy example of a pest population density function $f(x)$ defined by Eq. (19). The simulated sampled pest population densities $f_i$, $i = 1, \ldots, N$ (filled diamonds) are shown at locations (6) for the number $N = 5$ of sample units. (b) Convergence of the relative approximation error $E_{rel}$ (7) for estimates obtained using the numerical integration formula (2) with different weight choices. The relative error corresponding to the sample mean density estimate (3) is denoted $E_{rel}^{stat}$ (filled circles, solid line). $E_{rel}^{trap}$ (filled squares, solid line) represents the error of the estimates obtained on implementing the trapezoidal rule (4), and $E_{rel}^{Simp}$ (filled triangles, solid line) is the error for the Simpson's rule estimates (5). The accuracy tolerance (8) is set as $\tau = 0.2$ (dotted line). An estimate is considered to be sufficiently accurate when its relative error lies below this threshold.

monitoring problem means that the number of sample units $N$ that can be used is limited. As such, employing a method with a faster rate of convergence thus achieving the required level of accuracy (8) for a smaller value of $N$ seems preferable.

Let us consider an example to compare the convergence rate of estimates $I_a$ obtained *via* methods (3)–(5) for precise values of the density function $f(x)$. To calculate the relative error (7) we require the exact quantity $I$. Furthermore, in order to study the convergence rate of an estimate we also need to be able to evaluate each estimate over a series of increasingly refined grids of sample units. Consequently, for the purposes of this example we take the pest population density function to be mathematically defined by a function $f(x)$ as

$$f(x) = \sqrt{x^7} + \frac{1}{10} \quad x \in [0, 1] \tag{19}$$

which is shown in Fig. 1a. Integrating analytically we obtain $I = \frac{29}{90} \approx 0.3222$. A regular grid of sample unit locations is generated according to the formula (6) for a fixed value of $N$. We simulate the sampling procedure by evaluating the function (19) at the computed points $x_i$, $i = 1, \ldots, N$ to produce a discrete set of data $f_i$, $i = 1, \ldots, N$. An estimate $I_a$ is found by means of formula (2) with the selected set of weights, and the relative approximation error $E_{rel}$ is calculated according to (7). The convergence rate is established by computing $E_{rel}$ over a series of grids as defined by (6). The first grid is formed by fixing $N = N_1$ for some initial value $N_1$, where $N_1$ is odd. The subsequent grid is then generated by recomputing the number of grid nodes as $N_s = 2N_{s-1} - 1$, where $N_s$ is the number of nodes on the new grid and $N_{s-1}$ is the number of nodes on the previous grid. This process is repeated as many times as desired.

Convergence curves for the error of estimates formed using the weight choices (3)–(5) are plotted in Fig. 1b. The motivation for considering an estimate other than the sample mean population density is well illustrated by this graph. It can be seen that the relative error of an estimate formed using the trapezoidal rule (4) converges to zero at a faster rate than that of the sample mean (3). Meanwhile Simpson's rule (5) yields even faster convergence. To clarify the implication of this faster convergence rate let us fix the accuracy tolerance as $\tau = 0.2$. Whereas the sample mean density requires a grid of $N \approx 9$ sample units to satisfy the condition (8) that the relative error lies below the tolerance $\tau$, the trapezoidal rule requires a grid of only $N \approx 5$ units. Furthermore, Simpson's rule achieves the desired accuracy on the initial grid of only $N = 3$ sample units. In fact, for $N = 3$ we have $E_{rel}^{stat} \approx 0.4363$ and $E_{rel}^{Simp} \approx 0.0105$; the estimate formed by Simpson's rule is over forty times more accurate than that provided by the sample mean density.

Unlike in the above example the exact quantity $I$ is unlikely to be available in the real life pest monitoring problem. Thus, the accuracy cannot be assessed by considering the relative approximation error $E_{rel}$ as defined by (7). Instead, the usual way to conclude about an accuracy of an estimate formed by numerical integration is to consider the asymptotic error estimates. These are usually given in the form (*e.g.* see (Davis and Rabinowitz, 1975))

$$E(h) = Ch^k, \tag{20}$$

where $h$ is the fixed distance between the sample units and $h$ is required to be small in order to provide asymptotic convergence at the rate $k$. The constant $C$ and the convergence rate $k$ in (20) depend on the numerical integration method of choice. It can be shown that when the sample locations are defined by (6) that the convergence rate for the sample mean is $k = 1$, and it is well known that the convergence rate is $k = 2$ and $k = 4$ for the trapezoidal rule (4) and Simpson's rule (5), respectively (*e.g.* see Davis and Rabinowitz (1975)). From these error estimates, we typically expect that the higher the order of convergence of a method, the more accurate the estimates produced will be.

The above discussion demonstrates that choosing the weights of the formula (2) differently to those defined by the formula (3), the estimate often used in pest monitoring, has potential benefits. Firstly there is the potential for obtaining a more accurate estimate of the pest abundance $I_a$, and furthermore, a prescribed accuracy tolerance $\tau$ may be achieved for a smaller number $N$ of sample units. However, let us emphasise again that the above discussion is based on the assumption of precise data. Hence, in the next subsection we investigate asymptotic convergence of the mean error (16) calculated from noisy data.

### *4.2. Convergence of the evaluation error in the presence of noise*

In order to outline the potential effect noise could have on the accuracy of an estimate and the convergence rate of a numerical integration method we consider the behaviour of the mean $\mu(\tilde{E}_{rel})$ and the limits of the interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ to which $\tilde{E}_{rel}$ belongs with probability P(z). Once the estimate of the pest abundance based on measured data becomes sufficiently close to the true pest abundance $I$, the quantity $\tilde{E}_{\min}$ becomes zero. Since $z$ and $I$ are constant, the behaviour of the convergence of $\mu(\tilde{E}_{rel})$ and $\tilde{E}_{\max}$ is dictated by the quantities $I - I_a$ and $\sigma_{\tilde{I}}$. In other words, it is determined by the relationship between the exact pest abundance $I$ and the approximation $I_a$ formulated from exact values of the pest population density, as well as the uncertainty associated with the estimate $\tilde{I}$ formulated from measured pest densities. The convergence $I - I_a \to 0$ has been discussed in the previous sub-section and we now look at the convergence of the uncertainty $u(\tilde{I})$. Let us introduce the term $u_{max}$ such that

$$u_{\max} = \max_{i \in \{1,\ldots,N\}} u(f_i).$$

From the above and Eq. (15) it follows that

$$u(\tilde{I}) = \sqrt{\sum_{i=1}^{N} w_i^2 u^2(\tilde{f}_i)} \leq u_{\max} \sqrt{\sum_{i=1}^{N} w_i^2}.$$

Let us first consider the uncertainty associated with an estimate $\tilde{I}_{stat}$ formed by the sample mean density. Substituting the weights (3) into the above and recalling the distance between the traps is fixed as $h = 1/(N-1)$ we obtain

$$u(\tilde{I}_{stat}) \leq \frac{u_{\max}}{\sqrt{N}} = \left(\frac{1}{h} + 1\right)^{-(1/2)} u_{\max}.$$

For small $h$ we have

$$u_{\max}\left(\frac{1}{h} + 1\right)^{-(1/2)} \approx h^{1/2} u_{\max},$$

thus the convergence of the uncertainty $u(\tilde{I}_{stat})$ is of the order $k = 1/2$. Similar expressions can be found for the uncertainty associated with an estimate $\tilde{I}_{trap}$ formed by implementation of the trapezoidal rule (4)

$$u(\tilde{I}_{trap}) \leq h\left(\frac{1}{h} - \frac{1}{2}\right)^{1/2} u_{\max} \approx h^{1/2} u_{\max},$$

and likewise that associated with the estimate $\tilde{I}_{Simp}$ formulated from Simpson's rule (5)

$$u(\tilde{I}_{Simp}) \leq \frac{h}{3}\left(\frac{10}{h} - 2\right)^{1/2} u_{\max} \approx h^{1/2} u_{\max}.$$

Thus the orders of convergence of the uncertainty terms $u(\tilde{I}_{trap})$ and $u(\tilde{I}_{Simp})$ are, as for $u(\tilde{I}_{stat})$, also $k = 1/2$ provided the distance $h$ between traps is small.

In our previous study (Embleton and Petrovskaya, 2014) of the effect noise has on the accuracy of an estimate formed by the trapezoidal rule, we found that when the number of traps $N$ is large, the uncertainty $u(\tilde{I})$ associated with the estimate is the dominant contribution to the error of an estimate $\tilde{I}$. Therefore, in our present study of the three methods (3)–(5) we anticipate that as $N$ increases, the quantities $\mu(\tilde{E}_{rel})$ and $\tilde{E}_{\max}$ will converge at a rate of $k = 1/2$ in accordance with the behaviour of the uncertainty associated with the estimate $\tilde{I}$ as described above rather than the order $k$ as described by the error estimate (20). This important conclusion will be further illustrated in Section 6.

## 5. The accuracy on integration of noisy data: coarse grids

We now turn our attention to the ecologically relevant case of coarse grids of traps where data available for integration are sparse. It is a widespread situation in ecological monitoring that financial, ecological and other restrictions require the number of traps installed in an agricultural field to be relatively small (Mayor and Davies, 1976; Northing, 2009). For example, the number of traps installed over an agricultural field in the United Kingdom very rarely exceed a few dozen (Blackshaw, 1983; Ferguson et al., 2000; Holland et al., 1999), where a linear size of the field is typically of the order of a few hundred metres.

In our previous work a coarse grid was defined as a grid where the asymptotic error estimate (20) does not hold and we should expect poor accuracy of pest abundance evaluation. In many cases having a small number of traps installed in the field means that we have a coarse grid of traps. However, we cannot always provide a direct link between poor accuracy and a small number of traps. We have discussed in Petrovskaya and Embleton (2013, 2014), Petrovskaya and Petrovskii (2010), and Petrovskaya et al. (2012) that the definition of a coarse grid of traps should be based on the properties of the population density distribution rather than a number of traps installed in the field. In particular, grid coarseness is related to the degree of heterogeneity, highly aggregated density distributions being the most difficult case for pest abundance evaluation. It has been shown in Petrovskaya and Embleton (2013) and Petrovskaya and Petrovskii (2010) that an estimate of pest abundance can be very inaccurate when the total pest population size is evaluated from a strongly heterogeneous density pattern, while the same grid of traps will provide very good accuracy for another, quasi-homogeneous, density distribution.

Since ecologists and farmers often have to deal with pest insect density distributions that have a considerable degree of aggregation (Comins et al., 1992; Malchow et al., 2008; Okubo, 1986), the study of coarse grids becomes an important topic in integrated pest monitoring programmes. Our results in Petrovskaya and Embleton (2014), Petrovskaya and Petrovskii (2010), and Petrovskaya et al. (2012) have been obtained for the evaluation from exact data and we now need to further investigate the accuracy on coarse grids in the case that data are randomly perturbed.

On coarse grids, we cannot rely upon the convergence rate to conclude about the accuracy of the mean error associated with a selected method of numerical integration. Generally, the mean error and the bounds of the credible interval are determined by the following quantities:

1. Uncertainty associated with the estimate $\tilde{I}$ formulated from measured pest densities.
2. The error $E_{rel}$ obtained when the exact pest abundance $I$ and the approximation $I_a$ are formulated from exact values of the pest population density.

We begin our study of coarse grids by discussing the uncertainty in the estimate $\tilde{I}$. Below we compare the uncertainty in the estimate obtained when the trapezoidal rule (4) is employed with the uncertainty for the sample mean (3) and Simpson's rule (5) estimates on coarse grids.

From (15) it is clear that the uncertainty associated with an estimate formed from measured data $\tilde{I}$ will increase in magnitude as the magnitude of the weights of the numerical integration method increase. On comparing the weights of the sample mean density estimate (3), the trapezoidal rule (4), and Simpson's rule (5) it can be seen that none of these methods has uniformly larger weights than another. For example, we recall that the weights of the trapezoidal rule are $w_1 = w_N = h/2$ and $w_i = h, i = 2, \ldots N-1$ whereas for the sample mean density estimate they are uniformly

$w_i \equiv 1/N$. Thus, whilst the weights corresponding to the interior nodes are larger for the trapezoidal rule than the sample mean density estimate, the converse is true for those at the exterior nodes. Consequently, employing a method which by the asymptotic error estimate (20) is ordinarily considered more accurate, could in fact lead to a larger associated uncertainty. For instance the use of Simpson's rule (5) may result in a larger uncertainty in the estimate $\tilde{I}$ than that yielded by the trapezoidal rule (4). This occurs when the following condition is satisfied

$$\frac{h^2}{9}(u_1^2+u_N^2)+\frac{16h^2}{9}\sum_{i=1}^{(N-1)/2}u_{2i}^2+\frac{4h^2}{9}\sum_{i=2}^{(N-1)/2}u_{2i-1}^2>\frac{h^2}{4}(u_1^2+u_N^2)+h^2\sum_{i=2}^{N-1}u_i^2,$$

where $u_i \equiv u(f_i)$. The above can be expressed as

$$u_1^2+u_N^2<C_1\sum_{i=1}^{(N-1)/2}u_{2i}^2-C_2\sum_{i=2}^{(N-1)/2}u_{2i-1}^2, \qquad (21)$$

where the coefficients are $C_1 = 28/5$ and $C_2 = 4$.

Likewise, Simpson's rule could lead to a greater uncertainty associated with the estimate $\tilde{I}$ than that associated with the sample mean density (3). This will happen when we have

$$\frac{h^2}{9}(u_1^2+u_N^2)+\frac{16h^2}{9}\sum_{i=1}^{(N-1)/2}u_{2i}^2+\frac{4h^2}{9}\sum_{i=2}^{(N-1)/2}u_{2i-1}^2>\frac{1}{N^2}\sum_{i=1}^{N}u_i^2.$$

Using the fact that $h = 1/(N-1)$ and rearranging gives

$$u_1^2+u_N^2<C_3\sum_{i=1}^{(N-1)/2}u_{2i}^2-C_4\sum_{i=2}^{(N-1)/2}u_{2i-1}^2, \qquad (22)$$

where we have $C_3 = (7N^2+18N-9)/(8N^2-18N+9)$ and $C_4 = (5N^2-18N+9)/(8N^2-18N+9)$.

Finally, implementing the trapezoidal rule will give rise to a larger uncertainty than the sample mean density when the following condition is satisfied:

$$u_1^2+u_N^2<C_5\sum_{i=2}^{N-1}u_i^2, \qquad (23)$$

where $C_5 = (8N-4)/(3N^2-8N+4)$.

The conditions (21)–(23) can be used to decide which method is best to use on a coarse grid of traps in order to reduce the uncertainty of evaluation. Consider, for example, the condition (23) and let a very coarse grid of $N = 3$ traps be installed in the domain. The inequality (23) is then written as $u_1^2+u_3^2<C_5u_2^2$, where $C_5 = 20/7 \approx 3$. Hence, if the spatial pattern of a density distribution is such that the density is concentrated close to the domain centre (see a sketch of the density function shown in Fig. 2a), the inequality (23) holds. The uncertainty generated by the trapezoidal rule (4) is in this case larger than the uncertainty generated by the method (3). On the contrary, if the density is localised close to the boundaries (see Fig. 2b) then the trapezoidal rule yields a smaller uncertainty of evaluation. Similar analysis can be done for conditions (21) and (22).

It follows from the conditions (21)–(23) that the error in the pest abundance evaluation on coarse grids depends on the spatial pattern of the density function when noisy data are used for evaluation. This result is in line with the results of our previous work (Petrovskaya and Petrovskii, 2010; Petrovskaya et al., 2012) where the error of evaluation on exact data has been discussed. Thus we want to reiterate here a conclusion already made for the case of exact data in our work (Petrovskaya and Embleton, 2014). Namely, we conclude that the knowledge of spatial pattern of the pest insect density distribution is crucial when pest abundance is evaluated on coarse grids and any information about spatial pattern must be used to its fullest extent.

Another factor that makes an impact on the accuracy of evaluation from perturbed data is the error $E_{rel}$ calculated from exact data. As we have already mentioned the asymptotic error estimates (20) do not hold on coarse grids of traps and that impedes any theoretical discussion of the error (7). Meanwhile, a key result of the work (Embleton and Petrovskaya, 2014) was that when the number $N$ of traps is too small, the dominant contribution to the error of an estimate $\tilde{I}$ formed from measured data is the relationship between an estimate formed from exact data $I_a$ and the true pest abundance $I$. In other words, it was found that if for small $N$ the error $E_{rel}$ incurred by approximating the pest abundance by means of numerical integration on exact data is already significant, then the additional error caused by noise in the data has little relative impact. This conclusion in Embleton and Petrovskaya (2014) came from numerical experiments conducted for the trapezoidal method of numerical integration. Hence in the next section we design several ecologically meaningful test cases in order to investigate the contribution of the error $E_{rel}$ on coarse grids for various methods of numerical integration. For each of those test cases the accuracy on fine grids will also be investigated.

## 6. Numerical test cases

In this section we test the conclusions of the previous section for a variety of ecologically meaningful data. First, we outline how

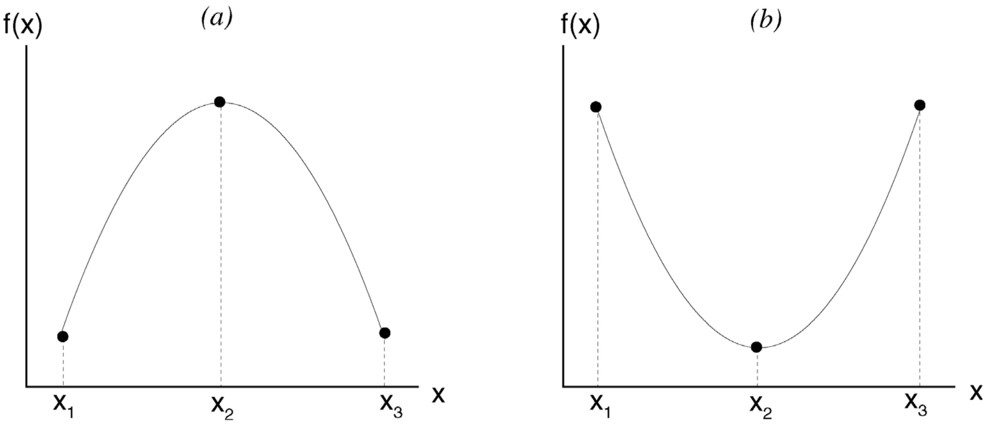


**Fig. 2.** A sketch of the density distribution $f(x)$ where evaluation of the pest abundance is done on a coarse grid of $N = 3$ traps. The measured data are available at the points $x_1$, $x_2$ and $x_3$ of a regular grid of traps. (a) The density is localised close to the centre of the domain. (b) The density is localised close to the boundaries.

we acquire such data. Then, estimates of the pest abundance are obtained by employing the methods (3)–(5) over a series of increasingly refined grids of traps, *i.e.* for increasing values of the number $N$ of traps. The mean error of estimates formed from measured data $\mu(\tilde{E}_{rel})$ is calculated for each value of $N$, as are the lower and upper limits of the $\alpha = 100P(z)$ percent credible interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ using (16)–(18), respectively. To assess the impact of noise on the accuracy of an estimate, a comparison is made with the relative errors of the estimates based on exact data $E_{rel}$. Plots of the convergence curves of all error quantities are given and the results are discussed.

### 6.1. Generating ecologically meaningful test cases

To form convergence curves of the quantities $E_{rel}, \mu(\tilde{E}_{rel}), \tilde{E}_{\min}$ and $\tilde{E}_{\max}$, we require the ability to form estimates on a series of increasingly refined grids of traps. Since it is difficult to obtain field data which satisfy these conditions, particularly in 1D, we instead choose to use computer simulated data to test the hypotheses from the previous section. The data are generated using the spatially explicit form of a predator prey model with the Allee effect (Murray, 1989; Turchin, 2003), which is given below in its dimensionless form:

$$\begin{aligned} \frac{\partial f(x,t)}{\partial t} &= d\frac{\partial^2 f}{\partial x^2} + f(1-f) - \frac{fg}{f+p}, \\ \frac{\partial g(x,t)}{\partial t} &= d\frac{\partial^2 g}{\partial x^2} + k\frac{fg}{f+p} - mg. \end{aligned} \tag{24}$$

The terms $f(x, t)$ and $g(x, t)$ represent the population density of the prey and predator, respectively, at the position in space $x$ at some time $t > 0$. The parameters in the system are as follows: $d$ is the diffusion coefficient, $p$ is the half-saturation prey density, $k$ is the food assimilation efficiency coefficient, and $m$ is the predator mortality. We consider the prey to be the pest insect. The pest population density $f(x, t)$ is found by numerically solving the above system of equations and considering a numerical solution at the fixed time $t = \hat{t} > 0$. Since the time is fixed we shall henceforth denote the pest population density as $f(x)$. The interested reader is referred to (Petrovskaya and Petrovskii, 2010) for the details of the non-dimensionalisation, as well as the numerical solving of the system of equations (24).

Since we intend to investigate the accuracy of numerical integration for a broad variety of density patterns, we generate six ecologically significant test cases from the model by inputting different parameter values. Plots of the resulting pest density functions are shown in Fig. 3a–f. The test cases are chosen such that the level of difficulty in obtaining an accurate estimate of pest abundance increases as we move from test case 1 through to test case 6. A test case is considered more difficult the higher the number $N$ of traps needed to obtain a sufficiently accurate estimate (*cf.* (Petrovskaya and Petrovskii, 2010; Petrovskaya et al., 2013)). Test case 1, as shown in Fig. 3a, is a smooth, monotonous function. The structure of the density function can therefore be detected from a small amount of data *i.e.* a small number $N$ of installed traps and an accurate estimate of abundance can be readily produced. The number of peaks present in the density function increases in the subsequent test cases until we reach test case 5 as shown in Fig. 3e which has a complicated multi-peak structure. More information about the pest density function, which corresponds to a higher number $N$ of installed traps, will be required to detect the more complex peak structure and thus obtain an accurate estimate. Meanwhile, test case 6 provides an example of the most difficult case whereby the pest population is located within a small subdomain of the field. The difficulties of handling such distributions, also known as *peak functions* have been discussed in detail in our previous works (Embleton and Petrovskaya, 2013; Petrovskaya and Embleton, 2013). If we consider a fixed number $N$ of installed traps, we expect the estimate of abundance to be most accurate for test case 1 and least accurate for test case 6.

### 6.2. Results and discussion

We now investigate the impact of noise on an estimate of pest abundance for the six ecologically meaningful test cases introduced in the previous subsection. The quantities $E_{rel}$, $\mu(\tilde{E}_{rel})$, $\tilde{E}_{\min}$ and $\tilde{E}_{\max}$, we need to assess the impact of noise on an estimate of pest abundance, all depend on the true value of the pest abundance $I$. In order to obtain $I$ the system of equations (24) was solved on a very fine regular grid as defined by (6). For all test cases shown in Fig. 3 the number of nodes on the fine grid was fixed as $N_f = 4097$. The data $f(x)$ obtained on the fine grid was then integrated using the trapezoidal rule (4) and the result was taken to be the 'exact' pest abundance $I$. The method (4) was selected as opposed to a more sophisticated method to limit the effect of round off error incurred on such a fine grid.

An estimate of pest abundance based on exact data $I_a$ is obtained using each of the weight choices (3)–(5), on a series of regular grids of traps and the error $E_{rel}$ is calculated as (7). The initial grid has the number of traps fixed as $N = N_1$, where $N_1$ is odd. The number of traps in subsequent grids is then calculated as $N_s = 2N_{s-1} - 1$ for $s \geq 2$. This process is repeated as many times as necessary to fully show the behaviour of the convergence. The quantities $\mu(\tilde{E}_{rel}), \tilde{E}_{\min}$ and $\tilde{E}_{\max}$ are then evaluated for each value of $N$ from (16)–(18) where the measurement tolerance has been set as $\nu = 0.3$. We have fixed $z = 3$ so the probability that a single realisation of the error $\tilde{E}_{rel}$ lies within the range $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ is $P(3) \approx 0.9973$.

Fig. 4a–f compare the error $E_{rel}$ of an estimate formed from exact data with the mean error $\mu(\tilde{E}_{rel})$ of an estimate formed from noisy data. The figures confirm that once the grid of traps becomes sufficiently refined, the convergence rate of the error quantity $E_{rel}$ behaves according to the asymptotic error estimates (20) and Simpson's rule (5) yields a more accurate estimate than the trapezoidal rule (4), which in turn is superior than the estimate provided by the sample mean density (3).

Meanwhile, it can also be seen from Fig. 4 that for each method of numerical integration the mean error of an estimate formed from noisy data, $\mu(\tilde{E}_{rel})$, converges at the slower rate of $k = 1/2$ as explained in Section 4. The difference in the convergence rates of $E_{rel}$ and $\mu(\tilde{E}_{rel})$ as shown in 4a—f demonstrates that when the number $N$ of traps is large, the accuracy of an estimate could be severely hampered by the presence of noise. In the presence of noise Simpson's rule (5) is not superior to the methods (3) and (4), as happens when precise data are used on fine grids. It should be noted, however, that as $N$ grows large the estimate of pest abundance based on exact data $I_a$ tends to the true pest abundance $I$ and hence the error of an estimate based on exact data $E_{rel}$ tends to zero. Consequently, the probability mass function of the quantity $\tilde{E}_{rel}$ transitions to a special case of the folded normal distribution, namely the half normal distribution (Daniel, 1959). In other words, the probability mass function skews towards zero, thus it becomes more probable that a smaller rather than a larger error will be obtained.

Fig. 5a–f show the mean errors $\mu(\tilde{E}_{rel})$ more clearly, as well as the quantity $\tilde{E}_{\max}$. The curves of $\tilde{E}_{\max}$ are shown to be parallel to that of $\mu(\tilde{E}_{rel})$ for larger $N$, therefore confirming that the convergence rate of $\tilde{E}_{\max}$ is also $k = 1/2$ as expected. Thus for large $N$, the uncertainty associated with the estimate of pest abundance $\tilde{I}$ caused by noise in the data $\tilde{f}_i, i = 1, \ldots, N$ is the dominant factor affecting the accuracy of an estimate. One interesting feature shown in Fig. 5a–f is that when the number

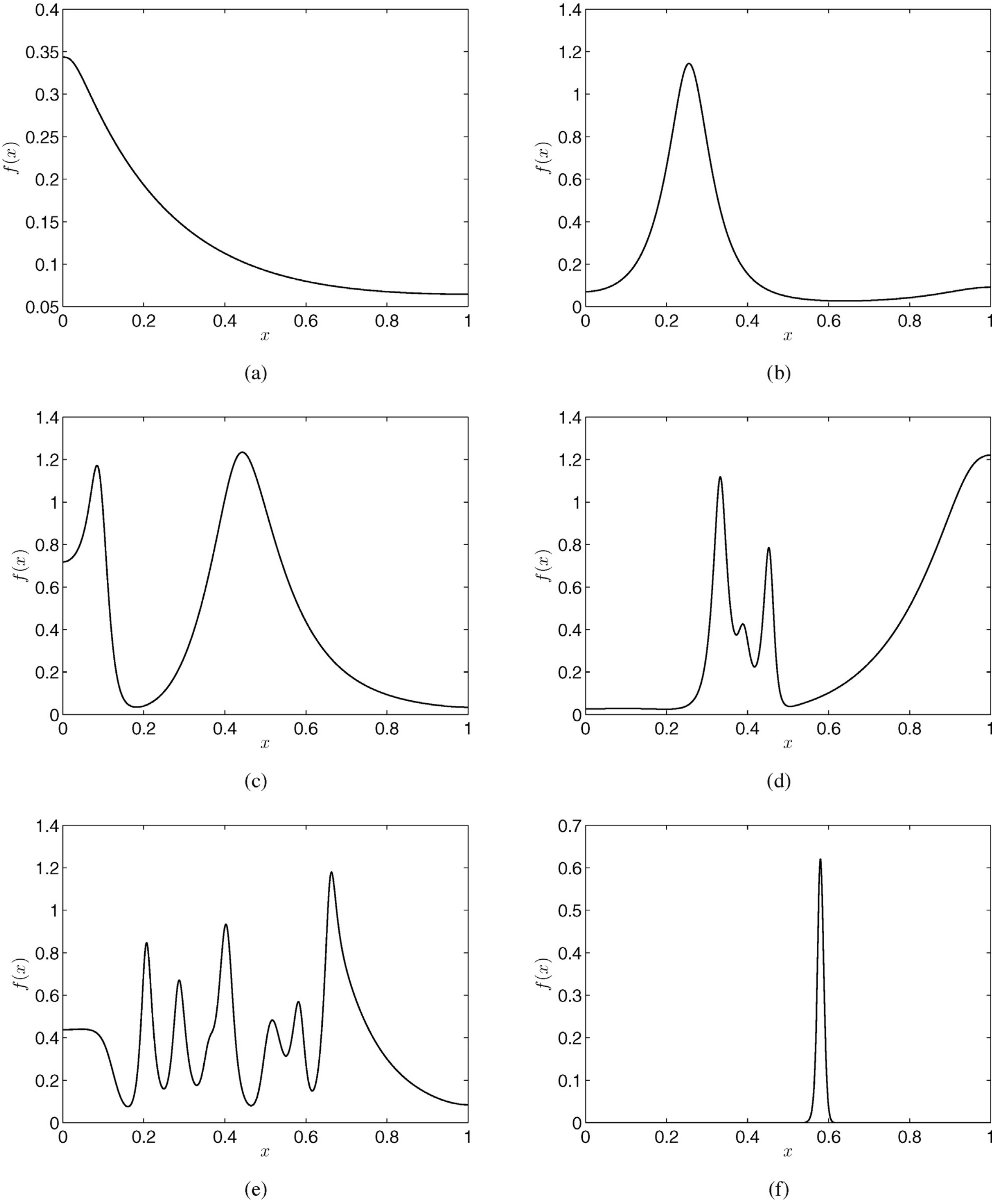


**Fig. 3.** Ecologically significant test cases. A spatial distribution *f*(*x*) of the pest population density function is obtained from the model (24) at different times *t* and for different values of the diffusivity *d*. (a) Test case 1, $d = 10^{-4}$, $t = 5$. (b) Test case 2, $d = 10^{-4}$, $t = 50$. (c) Test case 3, $d = 10^{-5}$, $t = 50$. (d) Test case 4, $d = 10^{-5}$, $t = 100$. (e) Test case 5, $d = 10^{-5}$, $t = 400$. (f) Test case 6, $d = 3 \times 10^{-6}$, $t = 10$. Other parameters $k = 2$, $h = 0.3$ and $m = 0.7$.

*N* of traps is large, the difference between the values of the quantity $\mu(\tilde{E}_{rel})$ for each of the methods of numerical integration (3)–(5) is very small, as is the difference between the values of $\tilde{E}_{\max}$. This confirms our previous conclusion that whilst the more sophisticated Simpson's method (5) outperforms the trapezoidal rule (4) and the sample mean density (3) as a means of estimating pest abundance on fine grids of traps when the data on the pest population density are precise, there is little difference between the methods when the data is noisy.

Table 1a–f gives the values of $\tilde{E}_{\min}$ for the test cases 1–6. As can be seen from (17), the definition of this quantity depends on the distance between the estimated pest abundance based on exact

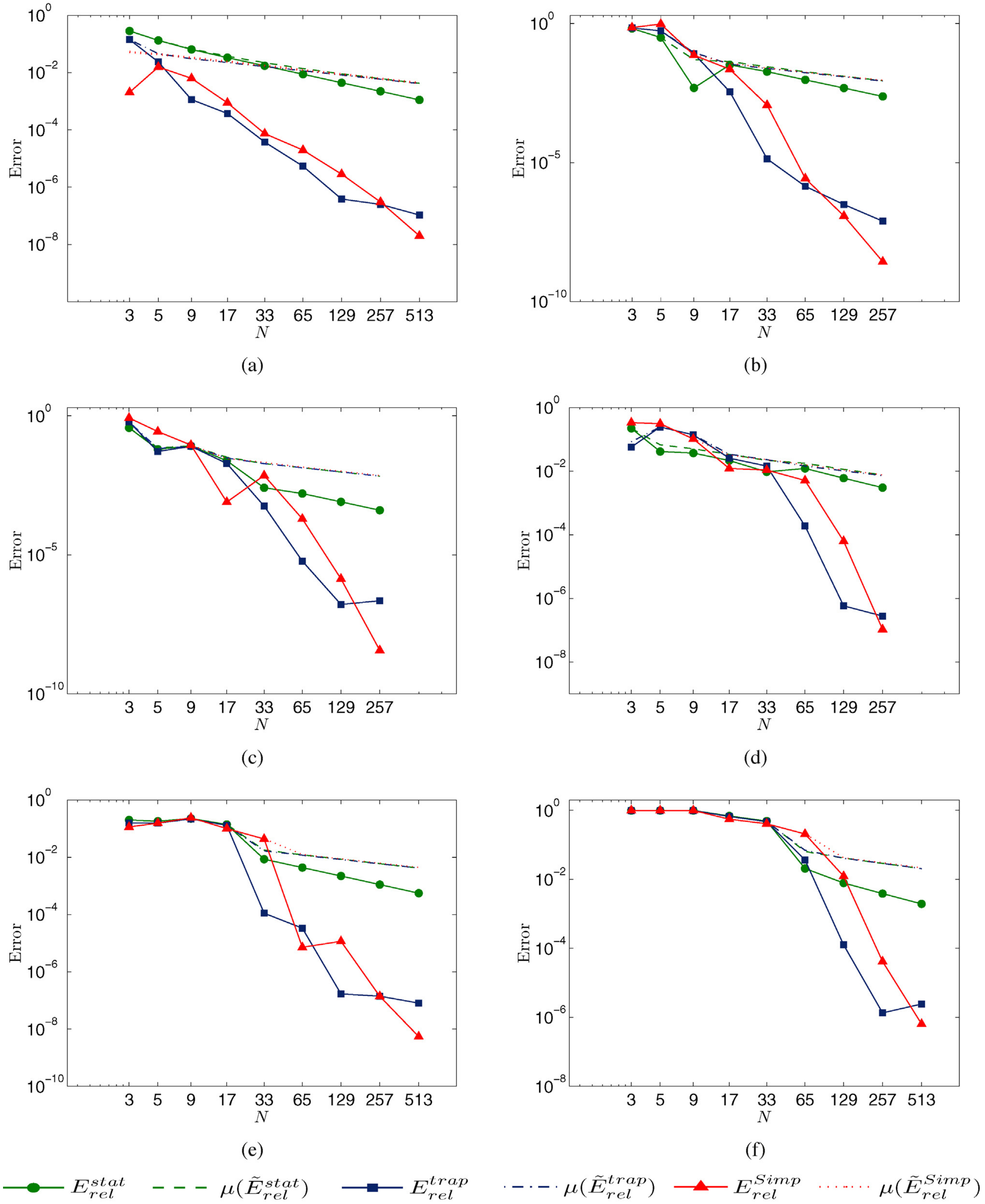


**Fig. 4.** Convergence curves for the density distributions $f(x)$ depicted in Fig. 3a–f, respectively. The mean error $\mu(\tilde{E}_{rel})$ of an estimate formed from noisy data (dashed lines) is compared with the error $E_{rel}$ of an estimate constructed from exact data (solid lines). The caption in all figures is as given above. The superscripts 'stat', 'trap' and 'Simp' indicate that the estimate of the pest abundance was calculated either as the mean density (3), or was formed using the trapezoidal rule (4) or Simpson's rule (5).

data $I_a$ and the true pest abundance $I$. A sufficiently accurate estimate $I_a$ is needed for the quantity $\tilde{E}_{\min}$ to be zero, thus the grid of $N$ traps needs to be sufficiently refined to resolve the heterogeneity of the pest population density. Table 1a–f confirms that the point at which $\tilde{E}_{\min}$ becomes consistently zero varies depending on the spatial pattern of the pest population density function of the corresponding test case (compare with Fig. 3a–f). For the easier to handle spatial density distributions *e.g.* the

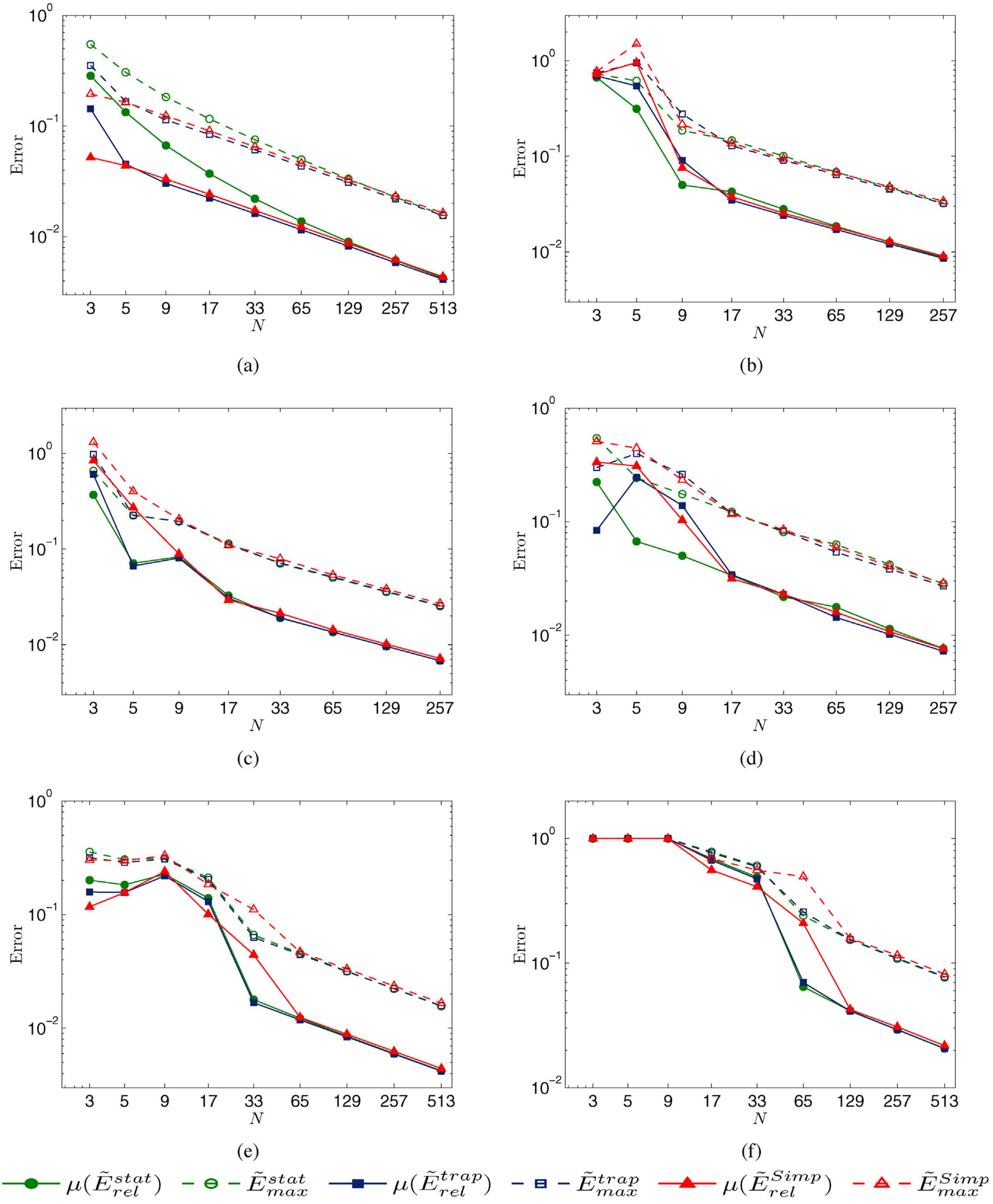


**Fig. 5.** Convergence curves for the density distributions *f*(*x*) depicted in Fig. 3a–f, respectively. Plots of the mean error quantity $\mu(\tilde{E}_{rel})$ are shown (solid lines) alongside the quantity $\tilde{E}_{\max}$ (dashed lines). The caption in all figures is as given above and the same superscript notation is used as in Fig. 4.

monotone function of test case 1 (see Fig. 3a), the quantity $\tilde{E}_{\min}$ is non-zero only for the estimate formed by the sample mean density (3) on the grid of $N = 3$ traps (see Table 1). Test cases 5 and 6 as shown in Fig. 3e and f on the other hand require further grid refinement before $\tilde{E}_{\min}$ becomes consistently zero. This happens for all numerical integration methods considered in the paper after the grid has been refined to $N = 33$ and $N = 65$ traps for test case 5 and test case 6, respectively.

Now let us consider the behaviour of the error quantities when the number $N$ of traps is small. Fig. 4a–f confirms the findings of our

**Table 1**
The quantity $\tilde{E}_{\min}$ for the test cases shown in Fig. 3a–f, respectively. The same superscript notation is used as in Fig. 4. For the larger values of $N$ where the values of $\tilde{E}_{\min}$ are not displayed, they are uniformly zero.

| $N$ | $\tilde{E}^{stat}_{\min}$ | $\tilde{E}^{trap}_{\min}$ | $\tilde{E}^{Simp}_{\min}$ |
|---|---|---|---|
| (a) | | | |
| 3 | 0.0057 | 0 | 0 |
| 5 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 |
| (b) | | | |
| 3 | 0.6088 | 0.6415 | 0.6686 |
| 5 | 0 | 0.1391 | 0.4180 |
| 9 | 0 | 0 | 0 |
| (c) | | | |
| 3 | 0.0819 | 0.2363 | 0.3724 |
| 5 | 0 | 0 | 0.1423 |
| 9 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 |
| (d) | | | |
| 3 | 0 | 0 | 0.1590 |
| 5 | 0 | 0 | 0.1742 |
| 9 | 0 | 0.1742 | 0 |
| 17 | 0 | 0 | 0 |
| (e) | | | |
| 3 | 0.0452 | 0 | 0 |
| 5 | 0.0600 | 0.0252 | 0.0118 |
| 9 | 0.1407 | 0.1308 | 0.1499 |
| 17 | 0.0679 | 0.0568 | 0.0170 |
| 33 | 0 | 0 | 0 |
| 65 | 0 | 0 | 0 |
| (f) | | | |
| 3 | 0.9982 | 0.9980 | 0.9977 |
| 5 | 0.9984 | 0.9984 | 0.9985 |
| 9 | 0.9985 | 0.9985 | 0.9985 |
| 17 | 0.5949 | 0.5697 | 0.4267 |
| 33 | 0.3757 | 0.3562 | 0.2627 |
| 65 | 0 | 0 | 0 |

earlier paper (Embleton and Petrovskaya, 2014). It can be seen from Fig. 4 that for small $N$ the accuracy of an estimate of pest abundance formed from noisy data is determined by the accuracy of an estimate based on exact data. That is, for small $N$ the quantity $\tilde{E}_{rel}$ is strongly dependent on $E_{rel}$ and is just slightly affected by the uncertainty caused by noise being present in the data. This is evident from the fact that in general the curves representing the mean error $\mu(\tilde{E}_{rel})$ of the estimates formed from noisy data lie close to their corresponding curves $E_{rel}$ when $N$ is small. In some cases the estimate of the pest abundance $I_a$ based on exact data already achieves good levels of accuracy even when $N$ is small. For example, for test case 1 (see Fig. 4a), this is evident for the estimates formed by implementing Simpson's rule. Here, there is a clear difference between the quantities $E^{Simp}_{rel}$ and $\mu(\tilde{E}^{Simp}_{rel})$ even on very coarse grids of $N = 3$ and $N = 5$ traps (compare the solid red line with closed triangles with the dashed red line). Whereas considering the estimates formed by the sample mean density (3) on the same coarse grids, it can be seen that there is little difference between $E^{stat}_{rel}$ and $\mu(\tilde{E}^{stat}_{rel})$ since the accuracy remains poor until the grid of traps is further refined (compare the solid green line with closed circles with the dashed green line).

Figs. 4f and 5f exhibit the behaviour on grids with a small number $N$ of traps whereby the quantities $E_{rel}$, $\mu(\tilde{E}_{rel})$, $\tilde{E}_{\max}$ and also $\tilde{E}_{\min}$ (see Table 1f) lie very close to each other. As discussed in Embleton and Petrovskaya (2014) this is the result of how we consider the noisy data $\tilde{f}_i$ to be related to the true population density values $f_i$. At the nodes of these grids the values of $f_i$ are very small. Since we essentially consider the $\tilde{f}_i$ to be a percentage of the corresponding $f_i$, in this instance the noisy data will lie close to the true data.

We continue to consider the ecologically relevant scenario where the number of grid nodes $N$ is small. Table 2a–f provides

**Table 2**
Contributions to the error of an estimate calculated from noisy data. For grids of a small number of traps and for each of the numerical integration methods (3)–(5), the quantities $u(\tilde{I})$, $E_{rel}$, and $\mu(\tilde{E}_{rel})$ are compared. The same superscript notation is used as in Fig. 4. For a fixed value of $N$, the greatest of each quantity is highlighted in bold and the lowest is given in grey text. The position of the bold/grey text in the $\mu(\tilde{E}_{rel})$ row matches that of the $E_{rel}$ row.

| $N$ | 3 | 5 | 9 |
|---|---|---|---|
| (a) | | | |
| $u(\tilde{I}^{stat})$ | **0.0121** | **0.0081** | **0.0055** |
| $u(\tilde{I}^{trap})$ | 0.0099 | 0.0067 | 0.0049 |
| $u(\tilde{I}^{Simp})$ | 0.0085 | 0.0068 | 0.0054 |
| $E^{stat}_{rel}$ | **0.2838** | **0.1327** | **0.0643** |
| $E^{trap}_{rel}$ | 0.1409 | 0.0235 | 0.0011 |
| $E^{Simp}_{rel}$ | 0.0021 | 0.0156 | 0.0063 |
| $\mu(\tilde{E}^{stat}_{rel})$ | **0.2839** | **0.1335** | **0.0668** |
| $\mu(\tilde{E}^{trap}_{rel})$ | 0.1428 | 0.0452 | 0.0304 |
| $\mu(\tilde{E}^{Simp}_{rel})$ | 0.0521 | 0.0439 | 0.0332 |
| (b) | | | |
| $u(\tilde{I}^{stat})$ | **0.0042** | 0.0229 | 0.0132 |
| $u(\tilde{I}^{trap})$ | 0.0037 | 0.0285 | **0.0148** |
| $u(\tilde{I}^{Simp})$ | 0.0036 | **0.0379** | 0.0108 |
| $E^{stat}_{rel}$ | 0.6681 | 0.3138 | 0.0048 |
| $E^{trap}_{rel}$ | 0.6948 | 0.5459 | **0.0823** |
| $E^{Simp}_{rel}$ | **0.7214** | **0.9595** | 0.0723 |
| $\mu(\tilde{E}^{stat}_{rel})$ | 0.6681 | 0.3139 | 0.0502 |
| $\mu(\tilde{E}^{trap}_{rel})$ | 0.6948 | 0.5459 | **0.0907** |
| $\mu(\tilde{E}^{Simp}_{rel})$ | **0.7214** | **0.9595** | 0.0760 |
| (c) | | | |
| $u(\tilde{I}^{stat})$ | 0.0403 | 0.0244 | 0.0170 |
| $u(\tilde{I}^{trap})$ | 0.0518 | **0.0263** | **0.0262** |
| $u(\tilde{I}^{Simp})$ | **0.0659** | 0.0182 | 0.0176 |
| $E^{stat}_{rel}$ | 0.3701 | 0.0628 | 0.0824 |
| $E^{trap}_{rel}$ | 0.6069 | 0.0526 | 0.0798 |
| $E^{Simp}_{rel}$ | **0.8438** | **0.2725** | **0.0888** |
| $\mu(\tilde{E}^{stat}_{rel})$ | 0.3701 | 0.0711 | 0.0830 |
| $\mu(\tilde{E}^{trap}_{rel})$ | 0.6069 | 0.0666 | 0.0806 |
| $\mu(\tilde{E}^{Simp}_{rel})$ | **0.8438** | **0.2725** | **0.0893** |
| (d) | | | |
| $u(\tilde{I}^{stat})$ | **0.0407** | **0.0255** | **0.0175** |
| $u(\tilde{I}^{trap})$ | 0.0306 | 0.0177 | 0.0145 |
| $u(\tilde{I}^{Simp})$ | 0.0205 | 0.0157 | 0.0164 |
| $E^{stat}_{rel}$ | 0.2205 | 0.0414 | 0.0369 |
| $E^{trap}_{rel}$ | 0.0568 | 0.2455 | **0.1384** |
| $E^{Simp}_{rel}$ | **0.3341** | **0.3084** | 0.1027 |
| $\mu(\tilde{E}^{stat}_{rel})$ | 0.2230 | 0.0670 | 0.0501 |
| $\mu(\tilde{E}^{trap}_{rel})$ | 0.0837 | 0.2455 | **0.1384** |
| $\mu(\tilde{E}^{Simp}_{rel})$ | **0.3341** | **0.3084** | 0.1031 |
| (e) | | | |
| $u(\tilde{I}^{stat})$ | 0.0190 | 0.0150 | 0.0105 |
| $u(\tilde{I}^{trap})$ | 0.0210 | 0.0160 | 0.0108 |
| $u(\tilde{I}^{Simp})$ | **0.0248** | **0.0176** | **0.0110** |
| $E^{stat}_{rel}$ | **0.2009** | **0.1828** | 0.2268 |
| $E^{trap}_{rel}$ | 0.1579 | 0.1567 | 0.2193 |
| $E^{Simp}_{rel}$ | 0.1148 | 0.1563 | **0.2401** |
| $\mu(\tilde{E}^{stat}_{rel})$ | **0.2009** | **0.1828** | 0.2268 |
| $\mu(\tilde{E}^{trap}_{rel})$ | 0.1580 | 0.1567 | 0.2193 |
| $\mu(\tilde{E}^{Simp}_{rel})$ | 0.1173 | 0.1563 | **0.2401** |

**Table 2** (*Continued*)

| $N$ | 3 | 5 | 9 |
|---|---|---|---|
| (f) | | | |
| $u(\tilde{I}^{stat})$ | **1.30e−06** | **9.53e−07** | **6.87e−06** |
| $u(\tilde{I}^{trap})$ | 1.54e−06 | 1.03e−06 | **7.13e−07** |
| $u(\tilde{I}^{Simp})$ | **1.89e−06** | **1.05e−06** | 7.41e−07 |
| $E_{rel}^{stat}$ | **0.99852** | 0.99867 | 0.998711 |
| $E_{rel}^{trap}$ | 0.99833 | **0.99861** | **0.99869** |
| $E_{rel}^{Simp}$ | **0.99815** | **0.99871** | **0.998714** |
| $\mu(\tilde{E}_{rel}^{stat})$ | **0.99852** | 0.99867 | 0.998711 |
| $\mu(\tilde{E}_{rel}^{trap})$ | 0.99833 | **0.99861** | **0.99869** |
| $\mu(\tilde{E}_{rel}^{Simp})$ | **0.99815** | **0.99871** | **0.998714** |

further evidence to support the assertion that for each numerical integration method, the magnitude of $\mu(\tilde{E}_{rel})$ is mainly defined by $E_{rel}$ on coarse grids of traps. In other words, the impact of the uncertainty associated with the estimate $u(\tilde{I})$ caused by noise in the density data is dominated by the error $E_{rel}$ which is imparted by the means of obtaining an estimate *i.e.* the method of numerical integration. Table 2a–f gives $u(\tilde{I})$, alongside the quantities $E_{rel}$ and $\mu(\tilde{E}_{rel})$ for all test cases on the grids of $N = 3, 5$ and 9 nodes. For each fixed value of $N$, the uncertainties $u(\tilde{I})$ associated with an estimate are compared for each numerical integration method. The maximum of these uncertainties is given in bold, and the minimum is given in grey text. The same comparison is made for the relative errors $E_{rel}$ of an estimate based on exact data, and the mean errors $\mu(\tilde{E}_{rel})$ of an estimate formulated from noisy data.

It is shown in Table 2a–f that the numerical integration method which yields the maximum or minimum value of $\mu(\tilde{E}_{rel})$ for a fixed number $N$ of traps is the same as that which generates the maximum or minimum value of $E_{rel}$ for all test cases. Therefore, the accuracy of an estimate based on exact data should be used to assess which method is superior when $N$ is small. The tables also demonstrate the point made in Section 5, that there are instances when the uncertainty associated with an estimate generated by Simpson's rule is greater than that which arises as a result of employing the trapezoidal rule and/or using the sample mean density to estimate pest abundance. For example, this occurs for test case 2 when $N = 5$ as shown in Table 2b as well as for test case 3 when $N = 3$. Other examples of this happening can be seen in the remaining tables, as can examples of when the uncertainty associated with an estimate formed by the trapezoidal rule exceeds that associated with the corresponding sample mean density estimate.

The accuracy control on coarse grids remains, perhaps, the most difficult issue in the general problem of pest abundance evaluation. Let us emphasise again that the asymptotic error estimates (20) (which are the conventional way of assessing a method of numerical integration) do not necessarily hold when $N$ is small. This is evident by inspection of the convergence curves in figures 4a–f and has been discussed at length in Embleton and Petrovskaya (2013), Petrovskaya and Embleton (2013), Petrovskaya et al. (2013), and Petrovskaya and Embleton (2013). As explained in Petrovskaya and Embleton (2013), on coarse grids the error can be considered a random variable and we have thus recommended that a method of numerical integration should instead be assessed probabilistically. In other words, the probability of obtaining a desired level of accuracy should be calculated rather than the error of an estimate. An initial methodology for such an assessment is presented in Petrovskaya and Embleton (2013) for exact data only, however the findings of this paper indicate that the results would also apply for noisy data as the effects of noise can be ignored on coarse grids.

## 7. Concluding remarks

We have considered the problem of pest abundance evaluation when data used for such evaluation have random error. Several methods of numerical integration employed in the evaluation problem have been compared in terms of their accuracy. For each method of numerical integration we have studied the behaviour of the mean error $\mu(\tilde{E}_{rel})$ arising when pest abundance is evaluated from randomly perturbed data. We have also investigated the credible interval $[\tilde{E}_{\min}, \tilde{E}_{\max}]$ to which the error of evaluation belongs with a given probability $P(z)$. Whilst the motivation for this work was to shed light on the problem of estimating pest insect abundance, the results could be applicable to abundance estimates for other kinds of species. This would be the case so long as the evaluation of abundance for that species requires the pest density, and that the results of the measurements made can indeed be converted to the density. Furthermore, it should be reasonable to assume that there is random error present in the density data and that it is normally distributed about the true density values.

In our study we have distinguished between evaluation with a large number of traps and a small number of traps, as different accuracy criteria should be applied in the former and latter case. If the number of traps is large (a fine grid of traps) the methods of numerical integration can be compared based on their convergence rate. The convergence of the mean error has been investigated for three different methods of numerical integration. It has been demonstrated in the paper that the mean error of an estimate formed from noisy data converges to zero at the same rate. This is despite the fact that the methods have different convergence rates when applied to exact data. The result of our paper confirm that for a large number of traps noise becomes a dominant feature of the approximation (*cf.* (Embleton and Petrovskaya, 2014)). This conclusion, however, does not immediately result in the recommendation to dismiss more advanced (and therefore more accurate on exact data) methods of pest abundance evaluation for the sake of methods less accurate yet easy to implement. It has been noted in the paper that it becomes more probable that a smaller rather than a larger error will be obtained in the theoretical limit when the number of traps $N \to \infty$. Hence a more accurate method of numerical integration can still be superior when perturbed data are considered on fine grids of traps. Further careful study of this topic is required and that will be a focus of our future research.

We have also studied an ecologically important case when the number $N$ of traps is small (a coarse grid of traps). On coarse grids, there is no convergence rate and another approach should be designed to compare methods of pest abundance evaluation. Generally, the mean error and the bounds of the credible interval are determined by the uncertainty associated with the estimate $\tilde{I}$ formulated from measured pest densities as well as the relationship between the exact pest abundance $I$ and the approximation $I_a$ formulated from exact values of the pest population density. We have shown that the uncertainty depends on the spatial pattern of the density function when two integration methods are compared on coarse grids. Hence, any *a priori* knowledge about the density distribution can be helpful in order to decide what integration method should be used to reduce uncertainty of the evaluation and obtaining such information will become a topic of our future work.

Meanwhile, it has been discussed in the paper that on coarse grids the most significant contribution to the error of an estimate $\tilde{I}$ formed from measured data is the relationship between an estimate formed from exact data $I_a$ and the true pest abundance $I$. Our numerical experiments confirmed that the impact of the uncertainty $u(\tilde{I})$ was negligible in many ecologically meaningful test cases where the difference $|I - I_a|$ was large. Thus on coarse grids of traps it is better to use a method that has a smaller error when exact data are considered and our further research will be focused on careful investigation of evaluation methods that can provide good accuracy on coarse grids.

Our study leaves several open questions, the issue of reliability being one of them. We have assumed in the paper that trap counts can be accurately converted into the values of the pest population density function and therefore our approach does not take into account errors related to such conversion. Meanwhile accurate interpretation of indirect measurements is considered as one of the most challenging issues in integrated pest monitoring, where a unified theory has not yet been developed despite various conversion techniques having been discussed in the literature (Arbogast and Mankin, 1999; Browde et al., 1992; Evans et al., 1983; Jansen, 1979; Petrovskii et al., 2012). Clearly, inaccurate conversion of trap counts into the pest population density distribution can significantly affect the results of pest abundance evaluation. This is especially true when trap counts are small, as any small change in a small trap count will result in a relatively big change in the value of the pest population density (Daniel, 1978). Hence estimating the conversion error and incorporating it into our approach to obtain a reliable estimate of the pest abundance is a challenging task that requires further careful investigation.

It should also be noted that the theoretical quantities $\mu(\tilde{E}_{rel})$, $\tilde{E}_{\min}$ and $\tilde{E}_{\max}$ used in this paper to assess the impact of noise on an estimate rely on the assumption that any measured pest density $\tilde{f}_i$ is normally distributed about the true pest density $f_i$ and belongs to the range (11) with probability $P(z)$. The counterpart to this assumption is that there is a chance that any $\tilde{f}_i$ can lie outside of this range. In particular the theory does not discount a measured pest density $\tilde{f}_i$ being negative. Of course a negative pest density is senseless, therefore instead each measured pest density should be considered to belong to a truncated normal distribution. The effects of such a truncation on the quantities $\mu(\tilde{E}_{rel})$, $\tilde{E}_{\min}$ and $\tilde{E}_{\max}$ will be the focus of future work.

Also, in the present paper we have considered uncorrelated noise making an implicit assumption that there is no interference between traps. In reality there may exist some correlation between trap counts in neighbouring traps, in particular when traps are installed sufficiently close to each other. Correlated noise may affect accuracy of pest abundance evaluation in a different way, and we therefore intend to investigate a topic of correlation in our future work.

Finally, another important direction of our research will be to investigate two-dimensional problems to extend our previous study of $2-d$ density distributions based on exact data. Our approach to pest abundance evaluation on randomly perturbed data can readily be applied to $2-d$ problems and our next goal is to implement various methods of numerical integration in problems where field data are available from real-life measurements.

## References

Arbogast, R.T., Mankin, R.W., 1999. Utility of spatial analysis in management of storage pests, 1519–1527. In: Zuxun, J., Quan, L., Yongsheng, L., Xianchang, T., Lianghua, G. (Eds.), Proceedings 7th International Working Conference on Stored Products Protection, 1998, vol. 2. Sichuan Publ., Sichuan, China.

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, 2008. Evaluation of Measurement Data Guide to the Expression of Uncertainty in Measurement, JCGM 100: 2008 GUM 1995 with Minor Corrections, first ed. Joint Committee for Guides in Metrology, Paris.

Blackshaw, R.P., 1983. The annual leatherjacket survey in Northern Ireland, 1965–1982, and some factors affecting populations. Plant Pathol. 32, 345–349.

Bolstad, W.M., 2007. Introduction to Bayesian Statistics. Wiley, NY.

Browde, J.A., Pedigo, L.P., Degooyer, T.A., et al., 1992. Comparison of sampling techniques for grasshoppers (Orthoptera: Acrididae) in soybean. J. Econ. Entomol. 85, 2270–2274.

Burn, A.J., Coaker, T.H., Jepson, P.C., 1987. Integrated Pest Management. Academic Press, New York, USA.

Byers, J.A., Anderbrant, O., Löfqvist, J., 1989. Effective attraction radius: a method for comparing species attractants and determining densities of flying insects. J. Chem. Ecol. 15, 749–765.

Comins, H.N., Hassell, M.P., May, R.M., 1992. The spatial dynamics of host-parasitoid systems. J. Anim. Ecol. 61, 735–748.

Cox, M.G., 2007. The area under a curve specified by measured values. Metrologia 44, 365–378.

Daniel, C., 1959. Use of half-normal plots in interpreting factorial two-level experiments. Technometrics 1, 311–341.

Daniel, W.W., 1978. Applied Nonparametric Statistics. Houghton Mifflin Company, Boston.

Davis, P.M., 1994. Statistics for describing populations. In: Pedigo, L.P., Buntin, G.D. (Eds.), Handbook of Sampling Methods for Arthropods in Agriculture. CRC Press, Boca Raton, USA, pp. 33–54.

Davis, P.J., Rabinowitz, P., 1975. Methods of Numerical Integration. Academic Press, New York, USA.

Embleton, N.L., Petrovskaya, N.B., 2013. On numerical uncertainty in evaluation of pest population size. Ecol. Complex. 14, 117–131.

Embleton, N.L., Petrovskaya, N.B., 2014. A novel approach to evaluation of pest insect abundance in the presence of noise. Bull. Math. Biol. 76, 718–743.

Evans, E.W., Rogers, R.A., Opfermann, D.J., 1983. Sampling grasshoppers (Orthoptera: Acrididae) on burned and unburned tallgrass prairie: night trapping vs. sweeping. Environ. Entomol. 12, 1449–1454.

Ferguson, A.W., Klukowski, Z., Walczak, B., Perry, J.N., Mugglestone, M.A., Clark, S.J., Williams, I.H., 2000. The spatio-temporal distribution of adult *Ceutorhynchus assimilis* in a crop of winter oilseed rape in relation to the distribution of their larvae and that of the parasitoid *Trichomalus perfectus*. Agric. Ecosyst. Environ. 95, 161–171.

Holland, J.M., Perry, J.N., Winder, L., 1999. The within-field spatial and temporal distribution of arthropods in winter wheat. Bull. Entomol. Res. 89, 499–513.

Jansen, M.J.W., 1979. How many victims will the pitfall make? Acta Biotheor. 28, 98–122.

Leone, F.C., Nelson, L.S., Nottingham, R.B., 1961. The folded normal distribution. Technometrics 3, 543–550.

Malchow, H., Petrovskii, S.V., Venturino, E., 2008. Spatiotemporal Patterns in Ecology and Epidemiology: Theory, Models, and Simulations. Chapman & Hall/CRC Press, London.

Mayor, J.G., Davies, M.H., 1976. A survey of leatherjacket populations in south-west England 1963-1974. Plant Pathol. 25, 121–128.

Metcalf, R.L., Luckmann, W.H. (Eds.), 1982. Introduction to Insect Pest Management. Wiley, London.

Murray, J.D., 1989. Mathematical Biology. Springer, Berlin.

Northing, P., 2009. Extensive field based aphid monitoring as an information tool for the UK seed potato industry. Asp. Appl. Biol. 94, 31–34.

Okubo, A., 1986. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. Adv. Biophys. 22, 1–94.

Pascual, M.A., Kareiva, P., 1996. Predicting the outcome of competition using experimental data: maximum likelihood and Bayesian approaches. Ecology 77, 337–349.

Pedigo, L.P., Rice, M.E., 2009. Entomology and Pest Management. Pearson Prentice Hall, NJ, USA.

Petrovskaya, N.B., Embleton, N.L., 2013. Evaluation of peak functions on ultra-coarse grids. Proc. R. Soc. A 469, http://dx.doi.org/10.1098/rspa.2012.0665.

Petrovskaya, N.B., Embleton, N.L., Petrovskii, S.V., 2013. Numerical study of pest population size at various diffusion rates. In: Lewis, M.A., Maini, P.K., Petrovskii, S.V. (Eds.), Dispersal, Individual Movement and Spatial Ecology, A Mathematical Perspective, Lecture Notes in Mathematics 2071. Springer-Verlag, Berlin Heidelberg, pp. 355–386.

Petrovskaya, N.B., Embleton, N.L., 2014. Computational methods for accurate evaluation of pest insect population size. In: Godoy, W.A.C., Ferreira, C.P. (Eds.), Ecological Modelling Applied to Entomology. Springer-Verlag, Berlin Heidelberg (in press).

Petrovskaya, N.B., Petrovskii, S.V., 2010. The coarse mesh problem in ecological monitoring.In: Proc. R. Soc. A, 466. pp. 2933–2953.

Petrovskaya, N.B., Petrovskii, S.V., Murchie, A.K., 2012. Challenges of ecological monitoring: estimating population abundance from sparse trap counts. J. R. Soc. Interface 9, 420–435.

Petrovskaya, N.B., Venturino, E., 2011. Numerical integration of sparsely sampled data. Simul. Model. Pract. Theory 19, 1860–1872.

Petrovskii, S., Bearup, D., Ahmed, D.A., Blackshaw, R., 2012. Estimating insect population density from trap counts. Ecol. Complex. 10, 69–82.

Raworth, D.A., Choi, M.-Y., 2001. Determining numbers of active carabid beetles per unit area from pitfall-trap data. Entomol. Exp. Appl. 98, 95–108.

Snedecor, G.W., Cochran, W.G., 1980. Statistical Methods. The Iowa State University Press, Ames.

Stern, V.M., 1973. Economic thresholds. Ann. Rev. Entomol. 18, 259–280.

Stern, V.M., Smith, R.F., van den Bosch, R., Hagen, K.S., 1959. The integration of chemical and biological control of the spotted alfalfa aphid. Part I: The integrated control concept. Hilgardia 29, 81–101.

Turchin, P., 2003. Complex Population Dynamics: A Theoretical/Empirical Synthesis. Princeton University Press, Princeton.