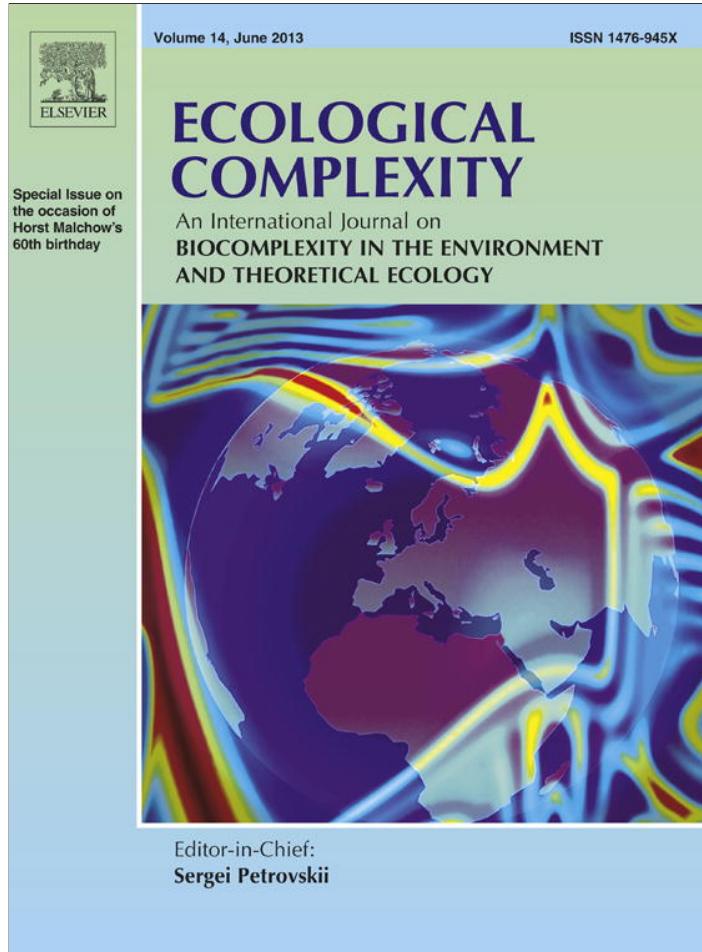


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at SciVerse ScienceDirect



Ecological Complexity

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

Original research article

On numerical uncertainty in evaluation of pest population size

N.L. Embleton*, N.B. Petrovskaya

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

ARTICLE INFO

Article history:

Received 5 July 2012

Received in revised form 26 November 2012

Accepted 29 November 2012

Available online 16 January 2013

Keywords:

Pest monitoring

Sparse data

Mean density

Single peak

ABSTRACT

Obtaining reliable estimates of pest insect species abundance is an essential part of ecological monitoring programs. It is often the case that data available for obtaining such estimates are sparse which in turn makes achieving an accurate evaluation difficult. This is especially true for strongly heterogeneous pest population density distributions. In our paper we discuss the accuracy of a mean density estimate when a certain class of high aggregation density distributions is considered and a standard statistical method is employed to handle sparse sampled data. It will be shown in the paper that conventional conclusions about the accuracy of the pest population size evaluation do not work when the data are sparse and a new approach is required. Namely, if the number of traps is small, an estimate of the mean density becomes a random variable with an error of high magnitude and we have to compute the probability of an accurate estimate rather than computing the estimate itself. We have obtained a probability of an accurate estimate based on the assumption that only one trap falls within a sub-domain where the pest population density is different from zero. The probability has been calculated for the one-dimensional and the two-dimensional problem.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Pest insects pose a serious problem to farmers as they can cause significant damage to crops. In recent decades careful monitoring of pest insects became a common agricultural procedure where monitoring techniques are being improved and new, more advanced methods of monitoring are being designed (Foster and Harris, 1997; Chapman et al., 2002; Mei et al., 2012). Consequently, the problem of pest population size evaluation is an essential part of the pest monitoring program. This problem received a lot of attention in the agricultural and ecological communities (Burn, 1987; Metcalf, 1982), as an estimate of the pest population size in an agricultural field is required to decide whether or not it is necessary to implement a control action (Ester and Rozen, 2005; Stern, 1973).

After an estimate of the pest population size has been obtained, the decision about any control action, often in the form of pesticides, is then made based on the relation of the estimated pest abundance to some 'critical density' (Binns et al., 2000) or 'action threshold' (Dent, 1991). In other words, the estimated pest abundance is compared to the limiting level of pests at which intervening is deemed to be worth the effort or expense. Making a correct decision is important as lack of action could lead to the loss

of valuable crops. It is also the case that a wrong decision could cause the unnecessary use of pesticides. Since pesticides are costly and can have detrimental effects on both the environment and human health, it is important to avoid their use when they are not needed (Blair and Zahm, 1991; Jepson and Thacker, 1990).

However making a correct decision about a control action remains a difficult and challenging problem. One aspect of this problem is uncertainty introduced by estimating the population size by a chosen evaluation technique rather than counting each individual pest insect in the field. Consider, for instance, the simple case of a single action threshold. If the average number of pests obtained from the evaluation procedure falls below the action threshold then the recommendation is that no action should be taken, whereas if the indication is that the average number of pests is greater than the action threshold then the decision should be to intervene (Stern, 1973). Obviously this decision making technique only becomes reliable if we have a reliable estimate that does not significantly differ from the true value of the pest population size. Hence the problem of accurate evaluation of the number of pest insects is pivotal in the decision making procedure.

Usually, an estimate is obtained through collecting samples and their subsequent analysis. One common method to collect samples of the pest insect species is to apply a trapping procedure, where a number of traps are installed across the monitored area. The traps are then emptied after a certain period of waiting, the insect species of interest is identified and the number of such insects in each trap is counted.

* Corresponding author. Tel.: +44 121 4142918; fax: +44 121 4143389.

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

It is worth noting here that more technically sophisticated methods such as mark-recapture (Lincoln, 1930) are also used by ecologists as they can provide a more accurate alternative to the trapping procedure above. However, for the routine monitoring of insect species methods of marking are time consuming, labour intensive and often too expensive (Hagler and Jackson, 2001) making such methods less practical.

Insect counts obtained via the trapping process provide the values of the pest population density at the trap locations (Byers et al., 1989; Raworth and Choi, 2001). An estimation of the pest abundance in the entire field can then be calculated from the information provided by the trap counts by statistical analysis (Davis, 1994; Seber, 1982), although some alternative methods have recently been developed (Petrovskaya and Petrovskii, 2010; Petrovskaya and Venturino, 2011; Petrovskaya et al., 2012). The accuracy of an estimate depends on the number N of traps installed in the field. In theory, the estimate must tend to the true pest population size when the number of traps gets infinitely large. In reality, the number of traps always remains finite and an approximate value always differs from the true pest population size. Moreover, it is in the interest of farmers and practical ecologists alike to make the number of traps as small as possible, as a large number of traps can bring considerable damage to the agricultural product. A large number of traps can also have a disruptive effect on the behaviour of the monitored insects and hence can result in a biased estimate of the population size. Another argument that must be taken into account is that a trapping procedure can be costly and labour-consuming. The reasons above impose significant restrictions on the number of traps installed over an agricultural field. It has been mentioned by many authors (e.g. see Blackshaw, 1983; Ferguson et al., 2003; Holland et al., 1999) that the number of traps installed in a typical agricultural field that has a linear size of a few hundred metres rarely exceeds a few dozen.

Let us define *the approximation error* as a difference between the true value and an estimated value of the pest population size. Given the restrictions in a routine monitoring procedure, an important question is to understand how big the approximation error can be if the sampled data are sparse (e.g. because the number of traps is small). This question has been the focus of ecological research for a long time (Dent, 1991; Vlug and Paul, 1986; Ward et al., 1985) and reliable recommendations have been provided on the minimum number of traps required for getting an accurate estimate of a particular pest insect species (Binns et al., 2000; Karandinos, 1976; Southwood and Henderson, 2000). Those recommendations, however, are based on the assumption that the pest insect population is spread over the monitored area with (almost) uniform density. While this assumption is true for many species, there also exists many ecologically important cases where the pest density is aggregated (Barclay, 1992; Ferguson et al., 2003; Petrovskaya et al., 2012). To our best knowledge, the cases of highly aggregated density distributions are not well studied in the literature, when a problem of pest abundance evaluation is considered. Hence in our paper we study an important case of an aggregated density distribution whereby the entire pest population is confined to a single sub-region (patch) within an agricultural field and the pest population is zero outside that patch. The term *highly aggregated* used throughout the paper refers to this particular type of density distribution. Such a distribution does have ecological significance as it corresponds to the beginning of the biological invasion (Shigesada and Kawasaki, 1997). Let us also note that the results obtained for a single patch can readily be extended to multi-patch density distributions.

The accuracy of population size estimation in the case of highly aggregated density was considered in our recent work

(Petrovskaya and Embleton, 2013) where we have used a method of numerical integration (Davis and Rabinowitz, 1975). The analysis there proved the existence of the threshold number $N_{threshold}$ of traps, where the desired accuracy of pest population size evaluation can be achieved for any $N \geq N_{threshold}$. In the case that the number N of traps is $N < N_{threshold}$, the approximation error becomes a random variable with a large magnitude because of the insufficient information about the sampled data. Consequently, we have to evaluate the probability p of achieving an approximation error smaller than a given tolerance rather than to evaluate the error itself (Petrovskaya et al., 2012; Petrovskaya and Embleton, 2013). An immediate consequence of this result is that the sampling plan has to be revisited when we have to deal with high aggregation density distributions, as an estimate of the pest population size per se becomes unreliable.

In the present paper we extend our earlier results to the case when a statistical counterpart of the numerical integration method is used to evaluate the pest population size for high aggregation density distributions. This statistical method is a space-implicit method based on computation of the arithmetic mean of the population density, where the average density is then multiplied with the area of a monitored field (Seber, 1982; Snedecor and Cochran, 1980). A widespread requirement that allows one to use a method of numerical integration is that the traps locations should be equidistant (Davis and Rabinowitz, 1975). Meanwhile the statistical approach does not specify trap locations and it implies that any distance between traps is admissible. Moreover, a common recommendation is to install traps randomly in the field in order to avoid biasing (Bliss, 1941; Legg and Moon, 1994; Reisen and Lothrop, 1999; Silver, 2008). It will be shown in the paper that, while the above recommendation works well for uniform density distributions (Perry, 1989; Alexander et al., 2005), random installation of traps may not always be optimal when highly aggregated density distributions are considered. We analyse the statistical method for an extreme case when only one trap falls inside the sub-domain of the non-zero density as a result of random installation of traps and conclude that an equidistant distribution of traps is better. Furthermore, it will be demonstrated in the paper that, if an equally spaced grid of traps cannot be provided, then installation of a certain (relatively small) number of traps will be the most beneficial from the accuracy viewpoint, as further increase in the number of traps will result in a less accurate estimate of the pest population size.

The paper is organised as follows. In Section 1 we discuss the problem statement and consider a few illuminating examples. It will be shown that the approximation error depends on how strongly heterogeneous a density distribution is across the monitored area. We then focus our attention on high aggregation density distributions in Section 2 where the problem is reduced to the one-dimensional case. It will be shown in Section 2 that the approximation error becomes a random variable if the number of traps is small and we explain how to compute the probability of accurate evaluation of the pest abundance. The results of Section 2 are illustrated by numerical examples in Section 3. We continue our study in Section 4 where the analysis is extended to the two-dimensional case. In Section 5 the results of our study are used to compare a random distribution of traps to a grid of equidistant traps. The conclusions and discussion of our results are provided in Section 6.

2. The problem statement: evaluation of the mean density when the data are sparse

In this section we discuss the accuracy problem when only a small number of traps are available in a pest population size

evaluation procedure. Let N traps be installed over an agricultural field. We shall assume that the number of insects caught in each trap is an accurate representation of the true population density in its catchment area. Whilst we are aware that this assumption is not entirely realistic, it is not within the scope of this paper to consider the problem of noisy data. However, it is worth mentioning here that the effectiveness of trapping has been the focus of much research (Legg and Moon, 1994; Petrovskii et al., 2012) and, with regard to our own work, two approaches can be considered that allow one to deal with inaccurate measurements of the pest population density. The first approach consists in developing a theoretical and methodological framework that enables a direct estimate of populations from trap counts. For instance, it has been shown in work (Petrovskii et al., 2012) that a 'mean-field' diffusion model is capable of revealing the generic relationship between trap catches and population density. The second approach is to take the noise into account when the data obtained as a result of trapping are handled by a method of numerical integration (Cox, 2007). Meanwhile we emphasize again that in the present paper we consider a number of insects in each trap as the true population density.

A statistical method commonly used in the evaluation of pest abundance depends on the sample mean pest population density (Davis, 1994). The mean density $M(N)$ is defined as

$$M(N) = \frac{1}{N} \sum_{i=1}^N u_i, \quad (1)$$

where u_i is the pest population density at each trap location (Snedecor and Cochran, 1980). The sample mean value $M(N)$ acts as an approximation to the true mean value E . An approximation S_{approx} to the actual pest population size S can then be found by multiplying the sample mean by the area of the field A , that is

$$S \approx S_{approx} = AM(N). \quad (2)$$

The decision about any control action with regard to pest insects is usually based on the assumption that the approximation S_{approx} is close to the true value S of the pest population size. While this assumption is correct in many cases, there can be density distributions where the formulae (1) and (2) present poor approximation of the true population size. Below we illustrate this statement by considering several density distributions over the unit square. Since the area $A = 1$, we will focus our attention on the estimate (1) where our aim is to demonstrate that the accuracy of approximation (1) depends strongly on the spatial heterogeneity of a density distribution.

Despite plenty of experimental data being available in the pest insect monitoring problem, in our further discussion we will use the data obtained as a result of computer simulation. This is because we need a sequence of values (1) obtained when we increase the number of traps for each consequent computation of (1). Hence we take our data from an ecologically sound mathematical model of population dynamics in order to be able to reconstruct the function $M(N)$ for various N of our choice. Namely, we consider the spatially explicit predator-prey model with the Allee effect (Murray, 1989; Turchin, 2003). In the dimensionless form the system is as follows:

$$\frac{\partial u(x, y, t)}{\partial t} = d \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + \beta u(u - b)(1 - u) - \frac{uv}{1 + \Lambda u}, \quad (3)$$

$$\frac{\partial v(x, y, t)}{\partial t} = d \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) + \frac{uv}{1 + \Lambda u} - mv, \quad (4)$$

where $u(x, y, t)$ and $v(x, y, t)$ are the densities of prey and predator, respectively, at time $t > 0$ and position (x, y) , d is the diffusion

coefficient, and the other parameters have evident meaning (Murray, 1989).

In order to obtain the population density distributions, the system ((3) and (4)) is solved numerically for a range of parameters, where the function $u(x, y, t)$ obtained as a result of computation is considered as the density of the pest insect in the problem. For the discussion of the numerical solution along with the choice of the initial and boundary conditions we refer the interested reader to our paper (Petrovskaya et al., 2012) where similar computer simulations have been carried out.

Let us fix the time t as $t = \hat{t} > 0$ and consider a snapshot $u(x, y) \equiv u(x, y, \hat{t})$ of a temporal-spatial density distribution $u(x, y, t)$. Numerical solution of (3) and (4) at any fixed time \hat{t} provides us with the discrete density distribution $u_i \equiv u(\mathbf{r}_i)$, $i = 1, \dots, N$ where grid nodes $\mathbf{r}_i = (x_i, y_i)$ are the points where traps are located. We then use the formula (1) to compute the mean value of a discrete function $u(x, y)$ considered on a grid of N traps.

Since the solution $u(x, y)$ is not available in closed form, we first solve the system ((3) and (4)) for a very big number of traps $N = N_f$, where $N_f = N_x N_y$ and $N_x = N_y = 2^{10} + 1$ is the number of traps installed in the x and y -directions of a rectangular grid, respectively. The mean value E computed for $N = N_f$ is considered as the true mean density and the spatial density distribution $u(x, y)$ obtained on a grid of N_f nodes at fixed time \hat{t} is stored for further computations. We then decrease the total number N of grid nodes and consider several approximations $M(N)$ of the true mean E obtained for smaller values of N . The formula (1) is used to obtain $M(N)$ every time that we decrease N in our computations. Let us note that we do not re-compute the density function $u(\mathbf{r}_i)$, $i = 1, \dots, N$ every time that a new number N is chosen. The values of $u(x, y)$ are always taken from the 'exact solution' computed on a grid of N_f nodes at time \hat{t} , where we make a projection of the function $u(\mathbf{r}_i)$, $i = 1, \dots, N_f$ obtained on the fine grid to a coarse grid every time that we take a new, smaller number N of nodes. For instance, if we consider the grid of N_f nodes and collect the data $u(\mathbf{r}_i)$ at each 4th grid node in the x -direction and in the y -direction alike, we obtain the density distribution on a coarser grid where the total number of nodes (traps) will be 16 times smaller in comparison with the original fine grid. The details of this computational technique are discussed in our previous work (Petrovskaya and Petrovskii, 2010).

Several density distributions obtained as a result of numerically solving (3) and (4) on a grid of N_f nodes are shown in Fig. 1a–c. The density functions $u(x, y)$ shown in Fig. 1a–c can be generated for a wide variety of the initial and boundary conditions; see Petrovskaya et al. (2012). For each density function $u(x, y)$ presented in Fig. 1a–c we show the true value E and the curve $M(N)$ in Fig. 1d–f respectively, where \tilde{N} is the number of nodes in each direction of a rectangular grid and the total number of nodes $N = \tilde{N}^2$. For the sake of convenience both $M(N)$ and E are scaled by the value of the true density, so that $E = 1$ in the figure.

It can be concluded from Fig. 1 that the accuracy of the estimation depends on how the pest insects are dispersed across the agricultural field. The density distribution $u(x, y)$ shown in Fig. 1a is a snapshot of the propagation of a continuous front. This density distribution is not strongly heterogeneous and the formula (1) gives us a very good estimate of the true mean density E even if the number of nodes is small (see Fig. 1d). Let us define the relative approximation error as $err = (|E - M(N)|)/E$, then the error is $err \approx 0.2$ even on a very coarse grid with $N = 9$ (the number of traps in each direction is $\tilde{N} = 3$).

The conclusion about accuracy on coarse grids is, however, different when the density distributions shown in Fig. 1b–c are considered. Both of these distributions present an ecologically important case of "patchy invasion" when the whole population is originally localised in a small sub-domain somewhere in the monitored area (Petrovskii et al., 2002, 2005). A strongly

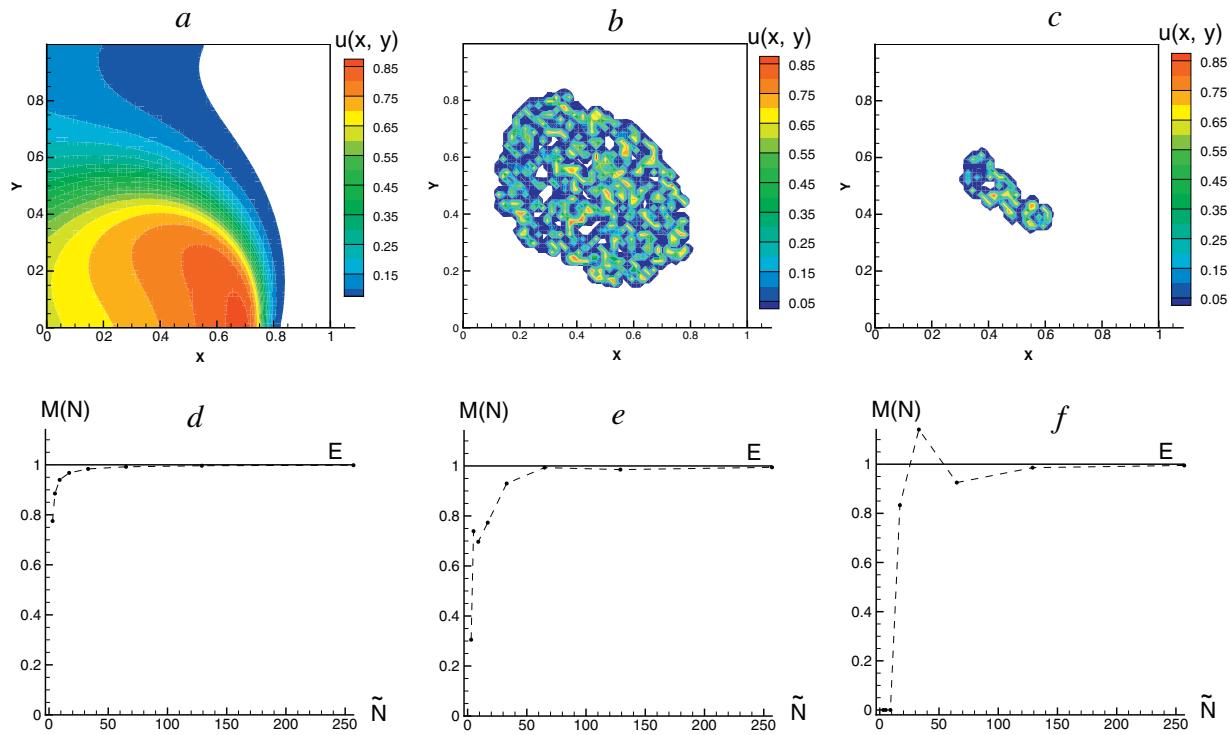


Fig. 1. (a–c) Pest population spatial density distributions $u(x, y)$ obtained from the population dynamics model ((3) and (4)). (d–f) The mean values $M(N)$ of the population density computed for density distributions (a–c), respectively, when the number N of traps is varied.

heterogeneous density pattern shown in Fig. 1b appears at the intermediate stage of patchy invasion. A small number of traps is not sufficient to resolve the sub-domain of the non-zero density and that is reflected by the accuracy of the formula (1) (see Fig. 1e). The accuracy on several coarse grids is below the value $err \approx 0.2$ that is achieved on the very first grid $N = 9$ in the case of the density distribution of Fig. 1a.

Finally, the risk of an inaccurate estimation is greatest when the pest population is located in a small area of the field, which corresponds to an early stage of the invasion (see Fig. 1c). It can be readily seen from Fig. 1f, that when a coarse grid of traps misses the entire sub-domain of the non-zero density, the accuracy of approximation (1) becomes unreliable.

One obvious yet important conclusion that can be derived from our consideration of Fig. 1c and f is that the accuracy of $M(N)$ depends on how many traps are stationed inside the sub-domain of non-zero density. Let us mention again that when high aggregation density distributions are considered in ecological problems, the exact location of the ‘peak sub-domain’ where the density is $u(x, y) \neq 0$ often cannot be predicted because it is affected by a broad variety of ecological and environmental factors. Hence, depending on the location of the non-zero density sub-domain, we can either miss it when a coarse grid of traps is installed, or install one or more traps within that sub-domain. Clearly in the former case an estimate of the mean density is $M(N) = 0$, while $M(N) > 0$ in the latter case and we have a better estimate of the mean population density.

Meanwhile the use of the statistical method (2) does not imply that the number of traps installed within the peak sub-domain always increases when we increase the total number of traps, as the method allows for random distribution of traps. As we already mentioned in the introduction, random (or pseudo-random) positioning of traps is a widespread technique widely recommended in the literature because it allows one to eliminate a bias-related error. For highly aggregated density distributions, however, installing traps in a random way means that detecting the patch of

pest insects is a matter of chance and the trade-off between the bias error and the approximation error may simply not exist. Below we consider this issue in more detail.

If there are N traps installed randomly across the field of the unit area $A = 1$, then the probability p_m that there are m traps within the peak sub-domain is defined by the binomial distribution (Spiegel, 1992):

$$p_m = \frac{N!}{m!(N-m)!} \delta^m (1-\delta)^{N-m}, \quad (5)$$

where δ is the area of the sub-domain within which the pest population is located and the following condition holds:

$$\sum_{m=0}^N p_m = 1. \quad (6)$$

Consider the probabilities p_0, p_1, \dots, p_N defined by the formula (5). Obviously, the first meaningful case is that we have one trap within the peak sub-domain as a result of a random installation of traps. Let us compute the probability $\hat{p}(N)$ given by the sum

$$\hat{p}(N) = p_0 + p_1 = (1-\delta)^N + N\delta(1-\delta)^{N-1}, \quad (7)$$

where p_0 is the probability that we have no traps within the peak sub-domain, and p_1 is the probability of having a single trap within that sub-domain. It follows from the expression (7) that the probability $\hat{p}(N)$ is dominant in the sum (6) when a narrow sub-domain δ is considered. In other words it is more probable that, for small δ the number of traps m located within the population of pests is either $m = 0$ or $m = 1$ than $m \geq 2$. This statement is illustrated in Fig. 2(a) where the probability $\hat{p}(N)$ has been plotted for an increasing number of traps N . The area δ for each graph shown in Fig. 2(a) has consequently been fixed as $\delta = 0.01, 0.06, 0.1, 0.2$, respectively. It can be seen from the figure that for the two smaller areas $\delta = 0.01$ and $\delta = 0.06$ the probability \hat{p} remains greater than 50% for the entire range of N displayed. For the larger choices of δ , that is the patch takes up 10 or 20% of the field, we

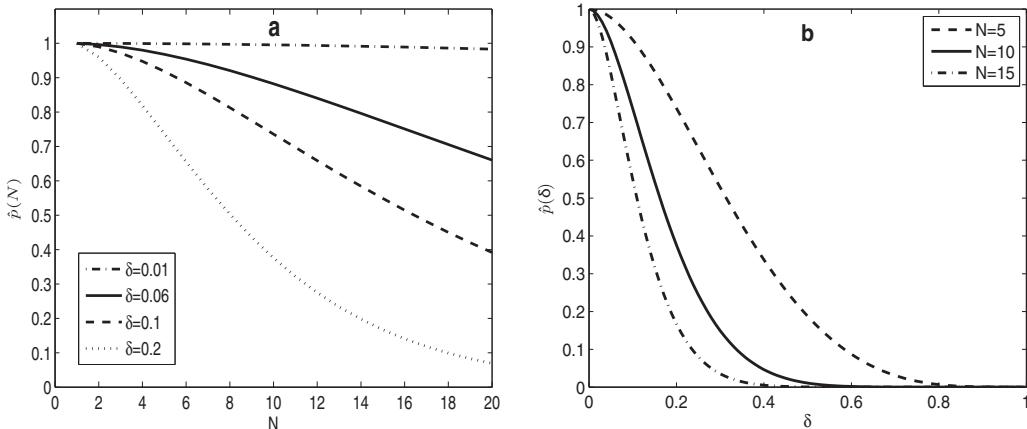


Fig. 2. The probability $\hat{p}(N)$ given by the formula (7). (a) The value of the total number of traps N varies and the area of the peak sub-domain δ is fixed. (b) N is fixed and δ varies.

have $\hat{p}(N) > 0.5$ for $N \leq 16$ and $N \leq 8$ respectively. In Fig. 2(b) a similar graph has been plotted, except this time the area of the patch of pests is varied for fixed values of the total number of traps N . In each case there is a range of sub-domain areas δ for which the condition $\hat{p}(N) > 0.5$ is satisfied. Thus the probability $\hat{p}(N)$ is dominant when the patch is small in comparison to the area of the field, a situation which corresponds to an early stage of biological invasion. The observations made above will be further discussed in Section 6.

The conclusion that we make here is that the bias problem is not very important when a single peak distribution is considered, as the most likely scenario is that we lose either all or all but one traps outside the peak sub-domain. Hence, the question we would like to investigate is whether a random trap distribution is still better than a grid of equidistant traps for highly aggregated density of pest insects. Clearly, a complete answer to this question would require investigation of all cases described by the formula (5), that is one trap within the peak sub-domain, two traps within the peak sub-domain, etc. However, in the present paper we restrict our discussion to the case of a single trap placed within the peak sub-domain even when the total number of traps is big. Despite this case not giving us a complete answer to the question of accurate estimation of the pest abundance, its study will allow us to identify and resolve several very important issues related to handling strongly localised density distributions. In particular, it will be revealed that a standard approach in the evaluation of the pest population size should be revisited when a highly aggregated density distribution is considered. In the next section we will show that the approximation error should be handled as a random variable and will compute the probability of obtaining an accurate estimate of the pest population size.

3. The analysis of a high aggregation density distributions: 1D case

3.1. Probability of obtaining an accurate estimate of the true mean population density

The spatial heterogeneity of a one-peak distribution $u(x)$ can be thought of as being constructed of two components – a peak region and a tail region. Clearly the peak is a dominant feature of the density $u(x)$ and it provides the main contribution to the mean density $M(N)$. Hence for the sake of our further discussion we formulate a simplified version of the distribution $u(x)$. In the peak region we consider $u(x)$ as a quadratic function of the width δ . Elsewhere we set the population density to be zero, therefore the

tail region is essentially ‘cut off’. We then have

$$u(x) \approx \begin{cases} g(x) = B - A(x - x^*)^2, & x \in [x_l, x_u], \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

where x^* is the location of the maximum of the peak, $A = -(1/2)(d^2u(x^*)/dx^2)$, $B = u(x^*)$. The peak width is $\delta = 2\sqrt{B/A}$ and the roots are $x_l = x^* - \delta/2$, $x_u = x^* + \delta/2$.

The above approximation of an arbitrary peak function $u(x)$ by a quadratic function $g(x)$ is very convenient, as it allows one to derive some important theoretical results about the accuracy of the formula (1). At the same time, the question of reliability arises when we replace our original density distribution $u(x)$ with a simpler function $g(x)$ as such a replacement introduces the error that depends on the peak width δ . However, the numerical study made in Petrovskaya and Embleton (2013) demonstrated that our conclusions about the quadratic approximation can be extended to the case when arbitrary peak functions are considered. Hence we consider the approximation as being reliable and use it in our further analysis.

Let N be the number of traps distributed over the domain and the density values $u(x_i) \equiv u_i$ be known at locations x_i . As we already mentioned in the previous section we intend to consider the limiting case of one trap installed in the domain $[x_l, x_u]$, where we have non-zero density $u(x)$. In other words, if the traps numeration is $i = 1, 2, \dots, N$, then we have $u_{i_0} \neq 0$ for fixed $i = i_0$ and $u_i = 0$ for any $i \neq i_0$. Let us re-define the index i_0 as $i_0 = 0$ for the sake of convenience. The trap location $x_{i_0} \equiv x_0$, where the density $u_{i_0} \equiv u_0 \neq 0$, is defined as

$$x_0 = x^* + \gamma \frac{\delta}{2},$$

where the parameter $\gamma \in [0, 1]$ as we only consider the right half of the peak sub-domain, because of the obvious symmetry of the peak (see Fig. 3b).

The density u_0 is computed as

$$u_0 \approx g(x_0) = B - A(x_0 - x^*)^2 = B(1 - \gamma^2), \quad \gamma \in [0, 1], \quad (9)$$

where we take into account that $\delta/2 = \sqrt{B/A}$.

Let now E be the true mean density and we require the approximation error to be small when E is approximated by $M(N)$:

$$\text{err} = \frac{|E - M(N)|}{E} < \tau, \quad (10)$$

where τ is a prescribed tolerance, $0 < \tau < 1$. Hence the mean value $M(N)$ computed when we use N traps should be within the range $(1 - \tau)E \leq M(N) \leq (1 + \tau)E$. (11)

We have already mentioned in the previous section that in ecological applications the exact location of the peak sub-domain cannot be predicted for a high aggregation density distribution. We now make an assumption that is crucial for our further discussion. Since the location of the sub-domain $[x_l, x_{ll}]$ is not known, with any location as likely to occur as another, the location x^* of the peak maximum can be considered to be a uniformly distributed random variable. Under the requirement that only one trap lies in the domain $[x_l, x_{ll}]$ the assumption about uniformly random location of the peak sub-domain can be re-formulated in terms of the location of the point x_0 . Namely, we fix the point x^* and then consider $\gamma \in [0, 1]$ as a uniformly distributed random variable in order to randomise the location of x_0 .

We now solve the inequalities (11) in order to see whether any location x_0 of a trap within the peak sub-domain can provide the desirable accuracy (11). The mean density is

$$M(N) = \frac{u_0}{N} = \frac{B(1 - \gamma^2)}{N}, \quad (12)$$

where (9) has been taken into account. Meanwhile, for a quadratic density distribution $g(x)$ the true mean density E is given by

$$E = \int_0^1 g(x)dx = \frac{2}{3}B\delta. \quad (13)$$

Substituting $M(N)$ and E in (11) we arrive at

$$(1 - \tau)B\hat{\delta} \leq \frac{B(1 - \gamma)^2}{N} \leq (1 + \tau)B\hat{\delta}, \quad (14)$$

where $\hat{\delta} = (2/3)\delta$.

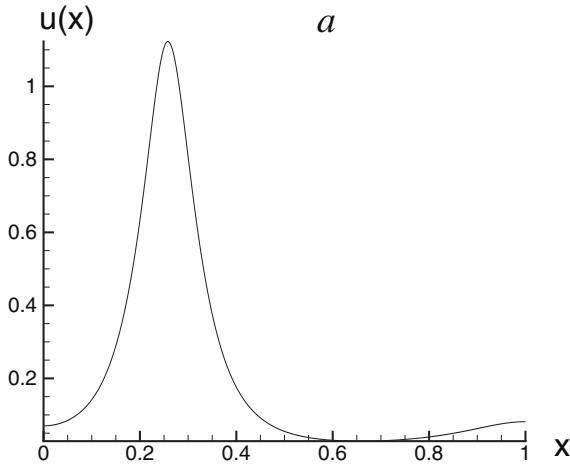
Consider the inequality

$$\frac{B(1 - \gamma)^2}{N} \leq (1 + \tau)B\hat{\delta}. \quad (15)$$

We have

$$1 - \gamma^2 \leq (1 + \tau)N\hat{\delta} \Rightarrow \gamma \geq \gamma_l = \sqrt{1 - (1 + \tau)N\hat{\delta}}, \quad (16)$$

where we have to choose a positive root, as $\gamma \in [0, 1]$. It follows from the solution of (16) that the number N of traps should be



restricted as

$$N \leq N^* = \frac{1}{\hat{\delta}(1 + \tau)}. \quad (17)$$

If the restriction (17) breaks, then the inequality (15) holds for any $\gamma \in [0, 1]$.

Let us now solve

$$\frac{B(1 - \gamma^2)}{N} \geq (1 - \tau)B\hat{\delta}. \quad (18)$$

Similar computation results in

$$1 - \gamma^2 \geq (1 - \tau)N\hat{\delta} \Rightarrow \gamma \leq \gamma_{ll} = \sqrt{1 - (1 - \tau)N\hat{\delta}}, \quad (19)$$

where the number of traps is restricted as

$$N \leq N^{**} = \frac{1}{\hat{\delta}(1 - \tau)}. \quad (20)$$

If the restriction (20) does not hold then the accuracy (11) is never achieved. In other words, if we have a big number $N > N^{**}$ of traps, but they are randomly distributed over the entire area $[0, 1]$, so that only one trap is positioned within the peak sub-domain, then the accuracy of the mean density evaluation will always be poor, as $M(N)$ will not be within the range (11).

Let us note that the number $N^* < N^{**}$ for any peak width δ and tolerance τ . Hence we have to consider the following cases

Case 1: $N \leq N^*$.

For any number of traps that is smaller than N^* the admissible range of trap locations x_0 where we can guarantee prescribed accuracy (11) is given by $\gamma_l \leq \gamma \leq \gamma_{ll}$. In other words, we require that $x_{0l} \leq x_0 \leq x_{0ll}$, where $x_{0l} = x^* + \gamma_l(\delta/2)$ and $x_{0ll} = x^* + \gamma_{ll}(\delta/2)$.

The same result holds when we consider a trap location at the left-hand side of the peak, $x_0 = x^* + \gamma(\delta/2)$, where $\gamma \in [-1, 0]$. We therefore have two subintervals $[-\gamma_{ll}, -\gamma_l]$ and $[\gamma_l, \gamma_{ll}]$ where the trap location within each of those subintervals will give us the accuracy required by (11). As the length of the entire interval is $\gamma \in [-1, 1]$ and a trap is randomly placed at any point of the peak sub-domain, then the probability $p(N)$ of obtaining a value $M(N)$

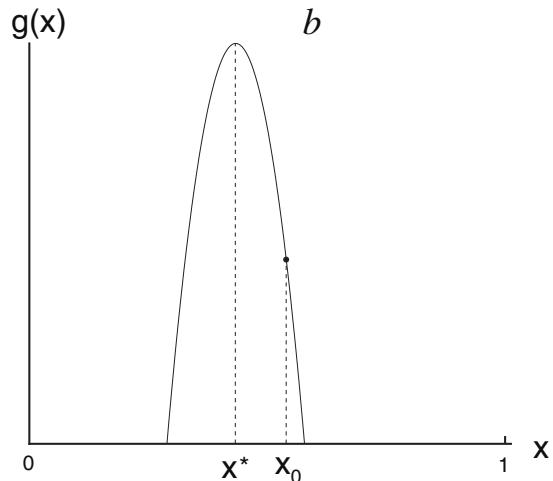


Fig. 3. (a) An example of high aggregation density distribution in a 1D ecological system. (b) Approximation of the one peak density distribution by a quadratic function. One trap is installed at the location x_0 within the peak sub-domain.

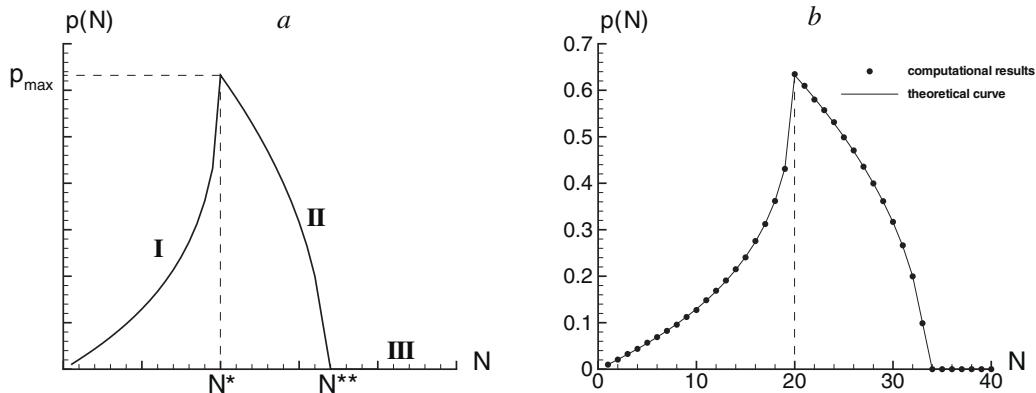


Fig. 4. The probability of obtaining an accurate estimate $M(N)$ of the true mean density E in the case when a single trap is installed within a peak subdomain formed by a quadratic function (8). (a) The theoretical curve. (b) Comparison of the theoretical curve and computational results. The probability is computed for the peak width $\delta = 0.06$ and the tolerance $\tau = 0.25$.

that meets the condition (11) is given by

$$p_I(N) = \frac{2(\gamma_{II} - \gamma_I)}{\gamma_{\max} - \gamma_{\min}}, \quad (21)$$

where $\gamma_{\min} = -1$, $\gamma_{\max} = 1$ and we multiply the range $\gamma_{II} - \gamma_I$ by 2 as we now consider the left-hand side and the right-hand side of the peak. Substituting γ_I and γ_{II} from (16) and (19) respectively in Eq. (21) we arrive at

$$p_I(N) = \sqrt{1 - (1 - \tau)N\hat{\delta}} - \sqrt{1 - (1 + \tau)N\hat{\delta}}. \quad (22)$$

The probability $p_I(N)$ for $N < N^*$ is shown as branch *I* of the graph in Fig. 4a. It can be seen from the graph as well as from the analytical expression (22) obtained for the probability $p(N)$ that the maximum value $p_{\max} = p_{\max}(\tau) = \sqrt{1 - ((1 - \tau)/(1 + \tau))}$ of the probability $p(N)$ is achieved when $N = N^*$. It is important to note here that the maximum probability is always $p_{\max} < 1$. At the same time the probability $p_{\max}(\tau)$ predictably grows when we make the tolerance τ bigger, that is $p_{\max} \rightarrow 1$ as $\tau \rightarrow 1$.

Case 2: $N^* < N \leq N^{**}$.

For any number of traps $N > N^*$ the inequality (15) always holds. Hence we only have the restriction (20) and the admissible range of γ becomes $\gamma \in [0, \gamma_{II}]$. The probability of obtaining an accurate estimate (11) is given by

$$p_{II}(N) = \sqrt{1 - (1 - \tau)N\hat{\delta}}. \quad (23)$$

The probability $p_{II}(N)$ defined for the number of traps $N^* < N \leq N^{**}$ is shown as curve *II* in Fig. 4a.

Case 3: $N > N^{**}$.

In the case that N is sufficiently big, the probability of the event that the error is within the range (11) is $p_{III}(N) = 0$ as we cannot meet the condition (18) (see branch *III* in Fig. 4a).

Let us note again that the results above are entirely based on the assumption that only one trap belongs to the peak sub-domain. However, we emphasize again that if a random distribution of traps over the domain is applied, then we cannot guarantee that more than one trap will be located in the peak sub-domain even when the total number N of traps is big (see discussion in Section 2). The branches *II* and *III* of the curve $p(N)$ in Fig. 4a will exist as long as we have a single trap within the peak sub-domain, no matter how big the total number N of traps is. Hence if we want to keep a random distribution of traps, our recommendation would be to restrict the number of traps as $N \approx N^*$ as this number of traps

would give us the biggest chance to get an accurate estimate of the mean density $M(N)$.

4. Numerical results

In this section the probability $p(N)$ will be obtained in several test cases by direct computation and compared with a theoretical curve obtained for a quadratic function (8). The first test case is to confirm that our theoretical results derived for a quadratic function are correct. Let us fix the peak width δ , the tolerance τ and the peak location x^* . We then consider the location x_0 of a trap as a random variable that is uniformly distributed over the interval $[x^*, x^* + \delta/2]$. In our computations we provide $n_r = 100,000$ realisations of the random variable x_0 for the fixed number N of traps, compute $M(N)$ and check the condition (11) for each realisation of x_0 . The probability $p(N)_{\text{comput}}$ of the accurate evaluation (11) of the mean density is then computed as

$$p(N)_{\text{comput}} = \frac{\hat{n}_r}{n_r}, \quad (24)$$

where \hat{n}_r is the number of realisations for which the condition (11) holds. We then increase the number of traps by one and repeat computation (24) for $N + 1$ traps. We stop increasing the number N , when the number N_L of traps becomes so big that the condition (20) breaks and we have $p(N_L) = 0$.

The probability $p(N)_{\text{comput}}$ of the accurate evaluation of the mean density is shown in Fig. 4b for the peak width $\delta = 0.06$ and the tolerance $\tau = 0.25$. We start from $N_1 = 1$ trap and then increase the number of traps until $N_L = 40$. It can be seen from the figure that all values of the probability $p(N)_{\text{comput}}$, $N = 1, \dots, 40$, computed by direct evaluation (24) belong to the theoretical curve $p(N)$.

The probability (22) and (23) is further illustrated for a quadratic function (8) in Fig. 5. Again, we assume that only one trap is located within the peak sub-domain and the location of that trap is random with respect to the position of the peak maximum. We make 100 random realisations n_r of the trap location x_0 and compute the error (10) for each realisation when the number of traps is fixed as $N = 10$. The integration error (10) computed for the function (8) is shown in Fig. 5a. The theoretical value of the probability $p(N)$ is $p(N) = 0.12$ when $N = 10$. This is well illustrated by the results shown in Fig. 5a where approximately 10% of the error values belong to the range (11). Clearly, the value $p \approx 0.1$ must tend to the theoretical probability $p(N) = 0.12$ when we increase the number of realisations n_r (cf. Fig. 4b). Consider now $N = N^*$, where the optimal number N^* of traps is defined from (17) as $N^* = 20$ for the peak width $\delta = 0.06$ and the tolerance $\tau = 0.25$.

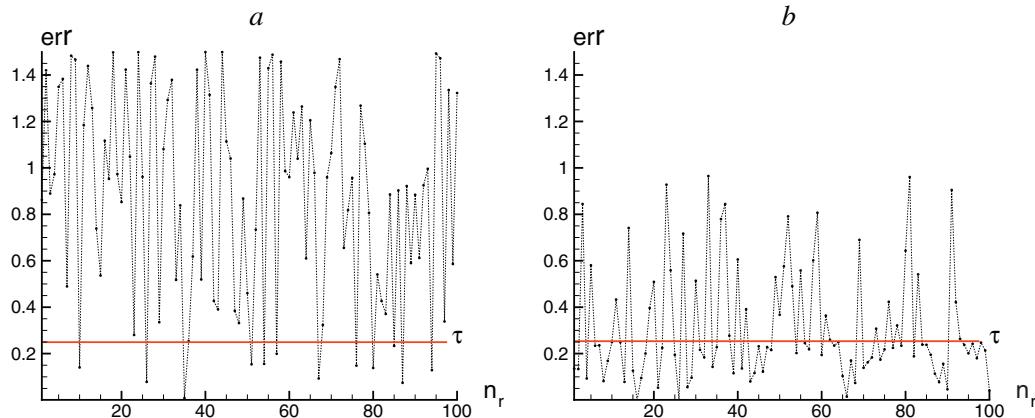


Fig. 5. Computation of the error (10) for a random peak location. The peak is given by a quadratic function (8) of width $\delta = 0.06$. The tolerance in the formula (10) is $\tau = 0.25$. A location of the peak is randomly generated 100 times and the error value is computed for each realisation $n_r = 1, 2, \dots, 100$ of the peak location. (a) The number of traps is $N = 10$. The probability of getting the error $err < \tau$ is low and most of the error values are beyond the required range. (b) The number of traps is $N = N' = 20$. The probability of getting an accurate answer $err < \tau$ achieves its maximum when $N = N'$ and most of the error values are within the required range.

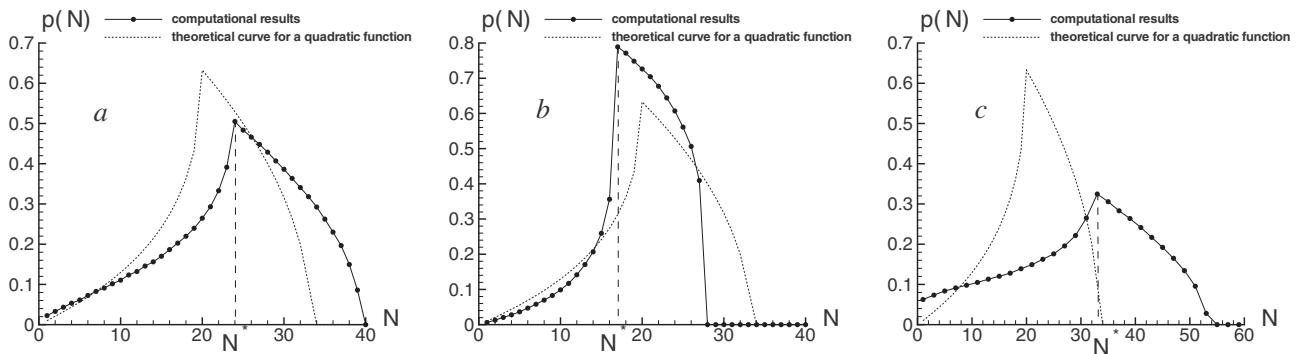


Fig. 6. Numerical test cases. The probability (24) (solid line) of an accurate answer (11) computed for (a) cubic function (25); (b) quartic function (26) and (c) normal distribution (27). For the functions (a)–(c) the peak width is chosen as $\delta = 0.06$. For each function (a)–(c) the probability (24) is compared with the theoretical curve obtained for a quadratic function (dashed line).

The probability of an accurate estimate is $p(N^*) = 0.66$ and this result is confirmed by the error distribution shown in Fig. 5b where most of the error values (every 2 out of 3) lie within the required range.

Let us now consider several standard test cases where high aggregation density distributions are different from the quadratic function (8). The test cases below are taken from the work (Petrovskaya and Embleton, 2013) where they have been investigated for the midpoint integration rule. Our first test case is to consider the cubic function

$$u(x) = \begin{cases} A(x - x^* + (\delta/3))(x - x^* - (2\delta/3))^2, & x \in [x^* - \delta/3, x^* + 2\delta/3], \\ 0, & \text{otherwise,} \end{cases} \quad (25)$$

where the peak width is $\delta = 0.06$ and $A = 30,000$. We apply the same computational procedure (24) as for the quadratic function discussed above to obtain the probability $p(N)_{\text{comput}}$ for various N . The probability graph for the function (25) is shown in Fig. 6a. Obviously, the probability graph obtained for a cubic function cannot coincide with the theoretical curve (22) and (23) (a dashed line in Fig. 6a). In particular, the critical number $N^* = 24$ is now different from the theoretical value $N^* = 20$ computed from (17) for $\tau = 0.25$ and $\delta = 0.06$. Nevertheless, it can be seen from the figure that the theoretical curve obtained for a quadratic function is a good approximation of the probability $p(N)$ computed for a cubic function (25).

A quartic function is defined as

$$u(x) = \begin{cases} A \left(\left(\frac{\delta}{2} \right)^4 - (x - x^*)^4 \right), & x \in [x^* - \frac{\delta}{2}, x^* + \frac{\delta}{2}], \\ 0, & \text{otherwise,} \end{cases} \quad (26)$$

where $A = 1,200,000$ and the peak width is again taken as $\delta = 0.06$. The probability graph for the function (26) is shown in Fig. 6b. It can be seen from the figure that the graph has a similar shape to the theoretical graph for the quadratic function, but the critical number $N^* = 17$ is again different from the number $N^* = 20$ obtained from the analysis of a quadratic distribution.

Finally, we consider a normal distribution

$$u(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{1}{2} \frac{(x - x^*)^2}{\sigma^2} \right), \quad (27)$$

that gives us an example of a peak function that is different from zero everywhere in the domain $x \in [0, 1]$. The peak width is defined by the parameter σ as $\delta = 6\sigma$ and we again consider $\delta = 0.06$. The probability graph computed from (24) for the function (27) is shown in Fig. 6c. It can be seen from the figure that the critical number $N^* = 33$ strongly differs from the number of traps obtained for a quadratic function with the same peak width. However, the shape of the graph is still similar to the theoretical curve (a dashed line in the figure) and the critical value N^* of traps provides the maximum probability $p(N^*)$. The presence of the critical value N^* in

each graph in Fig. 6 remains the most essential feature of our analysis.

4.1. Ecological test cases

In this subsection we verify our estimate (17) of the critical number N^* of traps when ecologically meaningful peak functions are considered. In our consideration of ecological test cases we use the pest density distributions $u(x)$ generated from the one-dimensional counterpart of the system (3) and (4). The parameters of the one-dimensional system of equations as well as the initial and boundary conditions required for its numerical solution are discussed in detail in the paper (Petrovskaya and Petrovskii, 2010). The function $u(x, t)$ considered at fixed time \hat{t} gives us a one-dimensional spatial density distribution $u(x, \hat{t}) \equiv u(x)$ of the pest insect over the unit interval $x \in [0, 1]$.

Similarly to the 2D case the properties of the spatial distribution $u(x)$ considered at a given time \hat{t} are determined by the diffusion d . Namely, the density distribution can evolve into the one-peak spatial pattern if the diffusion is $d \ll 1$. The examples of one-peak density distributions are shown in Fig. 7, where the functions $u(x)$ have been obtained from the numerical solution for the diffusion coefficient $d = 10^{-4}$ and $d = 10^{-5}$ respectively. The diffusion coefficient d is a controlling parameter in the problem and it has been discussed in Petrovskii et al. (2003), Malchow et al. (2008) that the peak width δ depends on the value of d . A simple estimate

of the function $\delta(d)$ can be written as

$$\delta = \omega\sqrt{d}, \quad (28)$$

see Petrovskii et al. (2003), where it has been argued that the coefficient ω in the expression (28) is relatively robust to changes in the parameter values, and can typically be considered as $\omega \approx 25$. Hence the critical number of traps can be evaluated as

$$N^* = \frac{1}{\delta(1 + \tau)} \approx \frac{C}{\sqrt{d}}, \quad (29)$$

where the coefficient $C(\tau) = 3/(2\omega(1 + \tau))$.

Consider the density distribution $u_1(x)$ shown in Fig. 7a. Since the diffusion coefficient is $d = 10^{-4}$, the estimate (29) gives us the number $N^* \approx 5$ for the tolerance $\tau = 0.25$. The probability graph obtained by direct computation is shown in Fig. 7b, where the number $N^* = 7$ taken from the graph is in good agreement with the theoretical estimate.

Let us now evaluate the number N^* in the case that we have the density distribution $u_2(x)$ shown in Fig. 7c. For the diffusion coefficient $d = 10^{-5}$ we have $N^* \approx 16$. The direct computation gives us $N^* = 23$ which is greater than the theoretical value of N^* obtained for a quadratic function. However, the results obtained for a quadratic function are still true for an ecologically meaningful density distribution. Namely, if traps are randomly installed over the monitored area and we cannot guarantee that more than one

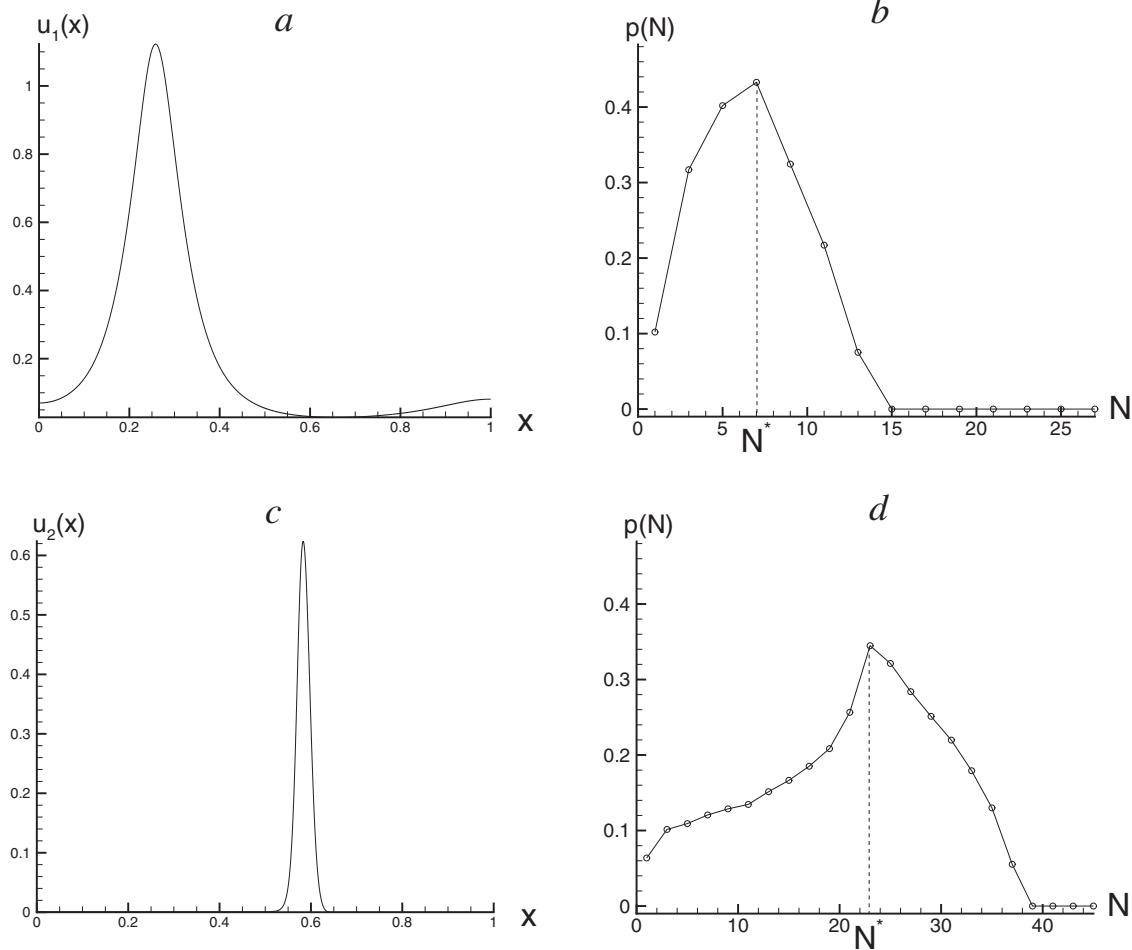


Fig. 7. Ecological test cases. (a) The spatial distribution $u_1(x)$ of the pest population density $u(x)$ for the diffusivity $d = 10^{-4}$. Other parameters along with the initial and boundary conditions used to generate one-dimensional density distributions are discussed in Petrovskaya and Petrovskii (2010). (b) The probability (24) of an accurate answer (11) computed for the density distribution $u_1(x)$ under the condition that a single trap is located within the peak sub-domain. (c) The pest population density $u_2(x)$ obtained for the diffusivity $d = 10^{-5}$. (d) The probability (24) computed for the density distribution $u_2(x)$.

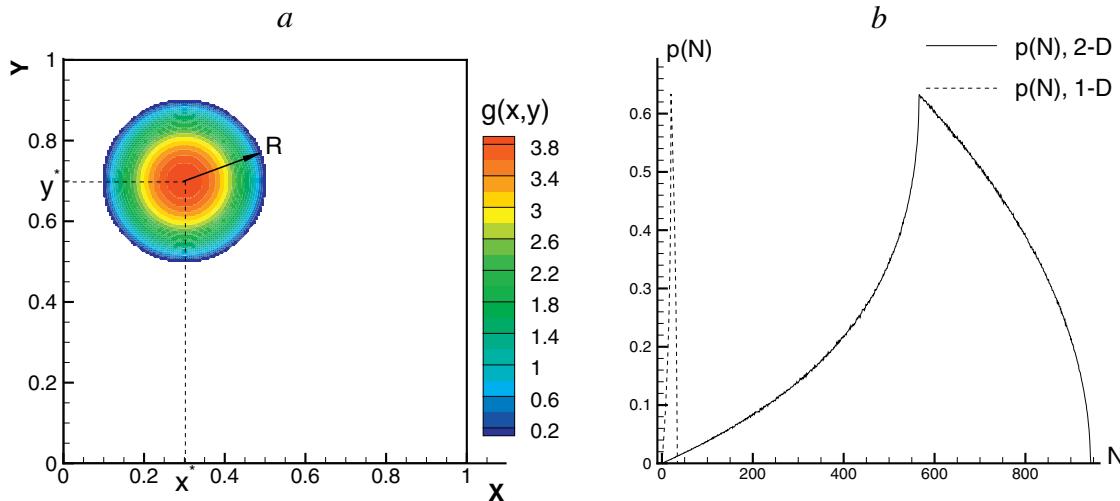


Fig. 8. (a) Quadratic distribution (30) where the location of the peak maximum is chosen to be $(x_0, y_0) = (0.3, 0.7)$ and the peak width is $\delta = 0.4$. (b) The probability curve $p(N)$ obtained for a 2D quadratic peak with the peak width $\delta = 0.06$. The probability curve for a 1D quadratic peak with the same peak width is shown as a dashed line.

trap will be installed within the peak sub-domain, then the best chance to get an accurate estimate of the mean density is when we use the number of traps close to the critical number N^* . Further increase in the number of traps reduces our chance of achieving an accurate answer.

5. The analysis of a high aggregation density distribution: 2D case

In this section we expand the results obtained from the analysis of the 1D problem to the more realistic 2D problem. We again focus on a high aggregation density distribution where there is a single peak in the domain. The domain of interest is now represented by the unit square $[0, 1] \times [0, 1]$. As in the analysis for the 1D problem, we consider the peak as a quadratic function, and ignore the tail region by setting the population density function $u(x, y)$ to be zero outside of the peak domain. That is we consider the population density function to be as follows:

$$u(x, y) \approx \begin{cases} g(x, y) = B - A((x - x^*)^2 + (y - y^*)^2), & (x, y) \in D_p, \\ 0, & \text{otherwise,} \end{cases} \quad (30)$$

where (x^*, y^*) is the location of the peak maximum. The peak sub-domain D_p is a circular disc of radius R , where $R = \sqrt{B/A}$ centred at (x^*, y^*) . This region can be seen in Fig. 8a. We define the peak width as $\delta = 2R$.

The probability analysis is similar to the 1D case and can be found in Appendix A. The probability $p(N)$ is shown in Fig. 8b, where the theoretical results (48) and (49) (solid line in the figure) are confirmed by direct computation. The same procedure used for the 1D case is applied in order to get computational results. Namely, the trap location (x_0, y_0) is randomised $n_r = 100,000$ times for a fixed number N , and the probability $p(N)_{\text{comput}}$ is calculated according to (24). The number of traps is then increased by one and the process is repeated. The peak width and the tolerance are fixed in computations as $\delta = 0.06$ and $\tau = 0.25$.

It can be seen from Fig. 8 that the shape of the graph $p(N)$ computed for a two-dimensional quadratic distribution is identical to the probability graph generated for a one-dimensional quadratic function (the dashed line in Fig. 8b; see also Fig. 4), except the critical number N_{2D}^* is different from the number of traps N^* obtained in the 1D case. That is a consequence of the definition (39)

where the function $u(x, y)$ is effectively a function of a single variable, $u(x, y) \equiv u(r)$. Thus both probability functions can be written in the following form:

$$p(N) = \begin{cases} \sqrt{1 - N(1 - \tau)\Delta} - \sqrt{1 - N(1 + \tau)\Delta}, & N \leq N^*(\Delta), \\ \sqrt{1 - N(1 - \tau)\Delta}, & N^*(\Delta) < N \leq N^{**}(\Delta), \\ 0 & N > N^{**}(\Delta), \end{cases} \quad (31)$$

where we now use a uniform notation $N^*(\Delta)$ and $N^{**}(\Delta)$ for the critical number of traps and the definition of the parameter Δ varies according to the number of dimensions in which we are working. In the 1D case we have

$$\Delta_{1D} = \frac{2\delta_{1D}}{3} \quad (32)$$

and in the 2D case

$$\Delta_{2D} = \frac{\pi R^2}{2} = \frac{\pi \delta_{2D}^2}{8}, \quad (33)$$

where δ_{1D} and δ_{2D} are the peak widths for the dimension denoted by the subscript.

It is clear that the theoretical probability curves will be the same when $\Delta_{1D} = \Delta_{2D}$. We can write the 1D peak width δ_{1D} in terms of δ_{2D} as

$$\delta_{1D} = \frac{3\pi\delta_{2D}^2}{16}. \quad (34)$$

Hence the probability of achieving an error (10) within a prescribed tolerance τ for a 2D peak can be calculated using the 1D theory. The critical number N_{2D}^* in (43) can then be computed as

$$N_{2D}^* = \frac{3}{(1 + \tau)2\delta_{1D}}. \quad (35)$$

For the 1D quadratic peak (8) with $\delta_{1D} = 0.06$ and the tolerance $\tau = 0.25$ we have that the number $N_{2D}^* = 566$ when a 2D counterpart with the same peak width $\delta_{2D} = 0.06$ is considered (see Fig. 8b). On the other hand, if we want to obtain the same critical number $N_{2D}^* = 20$ as in the 1D case, we have to set the peak width $\delta_{2D} = \sqrt{(16\delta_{1D}/3\pi)} = 0.3192$.

At the same time it is worth noting that the relation (34) between 1D and 2D problems is accurate for a quadratic function only. For a spatial distribution different from a quadratic function (8), Eq. (34) gives us an approximate estimate of the peak width

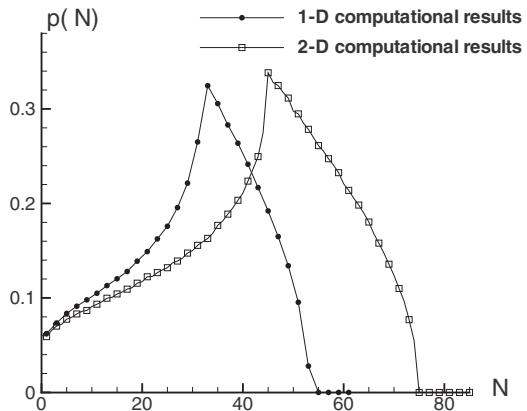


Fig. 9. Probability curves for the normal distributions (27) and (36) with the 'same' peak width $\delta_{1D} = 0.06$, $\delta_{2D} = 0.3192$.

and therefore an approximate value of the number N_{2D}^* of traps. Consider for example, a two-dimensional counterpart of the normal distribution (27). The function $u(x, y)$ is given by

$$u(x, y, t) = \frac{U_0}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x - \hat{x}_1)^2 + (y - \hat{y}_1)^2}{2\sigma^2}\right), \quad (36)$$

where the peak width is $\delta = 6\sigma$. Let us have $\delta = 0.06$ for a 1D distribution (27). It has been discussed above that the estimates (34) and (35) give us the peak width $\delta_{2D} = 0.3192$ for which the critical number N_{2D}^* in the 2D case should be the same as in the 1D case. The probability graphs for a 1D distribution (27) with $\delta = 0.06$ and a 2D distribution (36) with $\delta = 0.3192$ are shown in Fig. 9, where we expect the two graphs to be the same. However, it can be seen from Fig. 9 that the probability graph obtained for the normal distribution (36) is shifted from the graph $p(N)$ obtained for the 1D normal distribution (27).

We conclude this section by considering a simple yet ecologically meaningful example of a highly aggregated density distribution in 2D. Namely, we focus our attention on the pest population density distribution obtained from numerical solution of the Eqs. (3) and (4) as shown in Fig. 10. We consider a distribution $u_1(x, y)$ where the peak is wide, that is it takes up a large portion of the entire domain (see Fig. 10a). We also look at a distribution $u_2(x, y)$ for which the peak is restricted to a much smaller subdomain (see Fig. 10c). The distribution $u_2(x, y)$ was formed by placing the peak from $u_1(x, y)$ on a domain ten times larger in each direction. This is essentially the same as considering a peak with width δ ten times smaller than the original distribution.

In each case, the peak sub-domain is defined as the region in which the pest population density is such that $u(x, y) \geq 10^{-4}$. The region outside of D_p , i.e. the tail region is then ignored. Let (\tilde{x}, \tilde{y}) denote the points which belong to the peak sub-domain D_p . The width of the peak in the x and y directions, δ_x and δ_y , are calculated as $\delta_x = \max(\tilde{x}) - \min(\tilde{x})$, $\delta_y = \max(\tilde{y}) - \min(\tilde{y})$. We then define the peak width δ to be $\delta = \min(\delta_x, \delta_y)$. The distributions $u_1(x, y)$ and $u_2(x, y)$ were found to have peak widths of $\delta = 0.848541$ and $\delta = 0.0848541$ respectively.

An estimate of the point (x^*, y^*) is given by $x^* \approx (\max(\tilde{x}) + \min(\tilde{x}))/2$, $y^* \approx (\max(\tilde{y}) + \min(\tilde{y}))/2$. The random location (x_0, y_0) of a trap within the peak sub-domain is generated as

$$x_0 = r \cos \theta + x^*, \quad y_0 = r \sin \theta + y^*,$$

where $r \in [0, R]$ and $\theta \in [0, 2\pi]$ are uniformly distributed random variables. As before, we consider $n_r = 100,000$ realisations of the trap location.

We assume there is only one trap in the peak sub-domain D_p . In accordance with the procedure previously outlined, we now calculate $p(N)_{\text{comput}}$ for the population distributions $u_1(x, y)$ and $u_2(x, y)$. The results are shown in Fig. 10b and d respectively. It can be seen from the figure that the probability curves obtained for density distributions $u_1(x, y)$ and $u_2(x, y)$ differ from the graphs $p(N)$ computed for 1D ecologically meaningful density distributions (cf. Fig. 7). The difference can be explained by the fact that the functions $u_1(x, y)$ and $u_2(x, y)$ present the simplest case of a peak function when a highly aggregated density distribution is almost constant in the peak sub-domain. Hence the value of the mean density does not depend on a random location of the point (x_0, y_0) and the value of N in the expression (1) can be considered as a scaling coefficient. Nevertheless, this simple test case confirms our conclusions made in Section 2 that random installation of a big number of traps does not result in an accurate estimate of the mean population density, as we have $p(N) = 0$ for a big number N of traps in both cases (see Fig. 10b and d).

6. Random distribution of traps vs a grid of equidistant traps

The analysis made in the previous sections for the 1D and 2D cases revealed that there exists a critical number N^* of traps for which the probability of an accurate answer achieves its maximum value. The estimate of N^* , however, does not take into account the whole complexity of the problem when we have to deal with a random distribution of traps. First of all, let us note that the probability $p(N)$ should be scaled by the probability $p_1(N)$ of the event that exactly one trap is installed within the peak sub-domain. According to the formula (5) the probability $p_1(N)$ is calculated as

$$p_1(N) = N\delta(1 - \delta)^{N-1}. \quad (37)$$

The probability $\tilde{p}_1(N)$ of having the error with the given range (11) when a single trap falls into the peak sub-domain is then given by $\tilde{p}_1(N) = p_1(N)p(N)$. The functions $p_1(N)$ and $\tilde{p}_1(N)$ are shown in Fig. 11a and b, respectively. It can be seen from Fig. 11b that the resulting probability $\tilde{p}_1(N)$ is much smaller than $p(N)$. For a quadratic function with the peak width $\delta = 0.06$ and the tolerance $\tau = 0.25$ the critical number \tilde{N}^* for which the resulting probability $\tilde{p}(N^*)$ has its maximum is $\tilde{N}^* = 20$ and the probability is $\tilde{p}(N) \approx 0.23$.

On the other hand, it has been shown in Petrovskaya and Embleton (2013) that a uniform grid of equidistant traps over the unit interval $[0, 1]$ provides the desirable accuracy of the mean density evaluation with the probability $p(N) = 1$ when the distance between traps is $d = \alpha\delta$, where δ is the peak width in the one-dimensional problem and the parameter α depends on the tolerance τ only. In other words, if we use a grid of equidistant traps then the desired accuracy (10) will be achieved for any number of traps $N > N_{\text{threshold}} = 1 + 1/(\alpha\delta)$. For a quadratic function with the peak width $\delta = 0.06$ and the tolerance $\tau = 0.25$ the threshold number providing the error below the given tolerance has been computed as $N_{\text{threshold}} = 21$ (i.e. the distance between traps is $d = 0.05$). Any equidistant grid of traps with the number $N > 21$ will then give us an accurate estimate of the pest abundance.

The discussion above can be summarised as follows. For the given value $\delta = 0.06$ of the peak width, on a grid of randomly distributed traps the probability of getting an accurate estimate increases with the number of traps and reaches the value $\tilde{p}(N) \approx 0.23$ for $N = 20$. However, on a regular grid of equidistantly placed traps with about the same number $N = 21$, an accurate result is obtained with the probability being equal to one. Therefore, for $N \sim 20$ (or larger), on a regular grid the estimate of the population

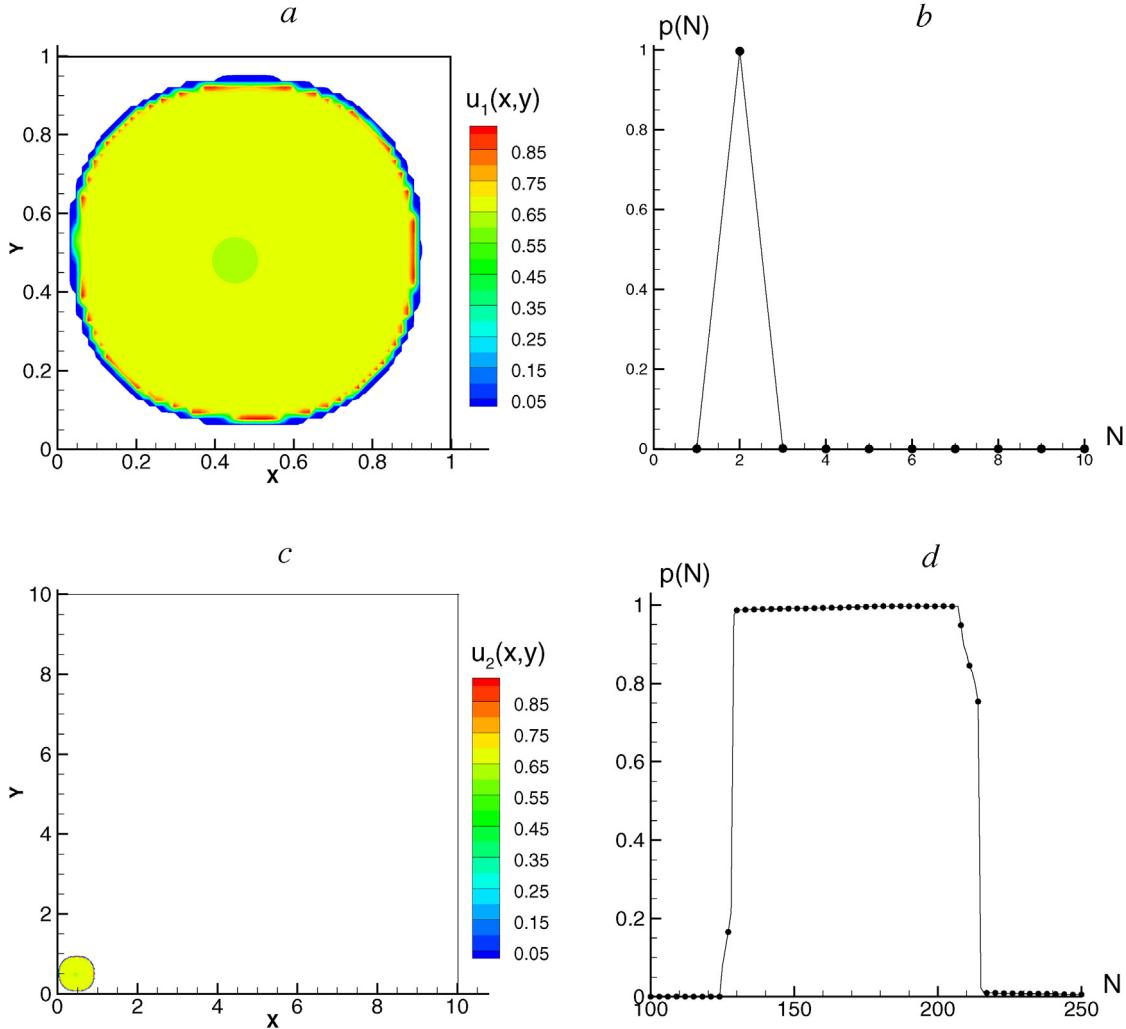


Fig. 10. 2D ecological test cases. (a) The spatial distribution $u_1(x,y)$ of the pest population density as predicted by the model ((3) and (4)) in the unit square. (b) The probability (24) of an accurate answer (11) computed for the density distribution $u_1(x,y)$ under the condition that a single trap is located within the peak sub-domain. (c) The pest population density $u_2(x,y)$ considered in the domain $D : x \in [0, 10], y \in [0, 10]$. (d) The probability (24) computed for the density distribution $u_2(x,y)$.

size is not a stochastic variable anymore, while on a random grid of the same size it still essentially stochastic and the probability of obtaining an accurate result remains relatively low. Thus a grid of equidistant traps is clearly more favorable compared to a grid of randomly distributed traps.

Clearly, the argument above is not complete, as for a random distribution of traps there is the possibility that more than one trap falls inside the peak sub-domain. Generally speaking, we have to make a similar computation for the probability $\tilde{p}_2, \tilde{p}_3, \dots, \tilde{p}_N$ in order to be able to conclude about the efficiency of a random

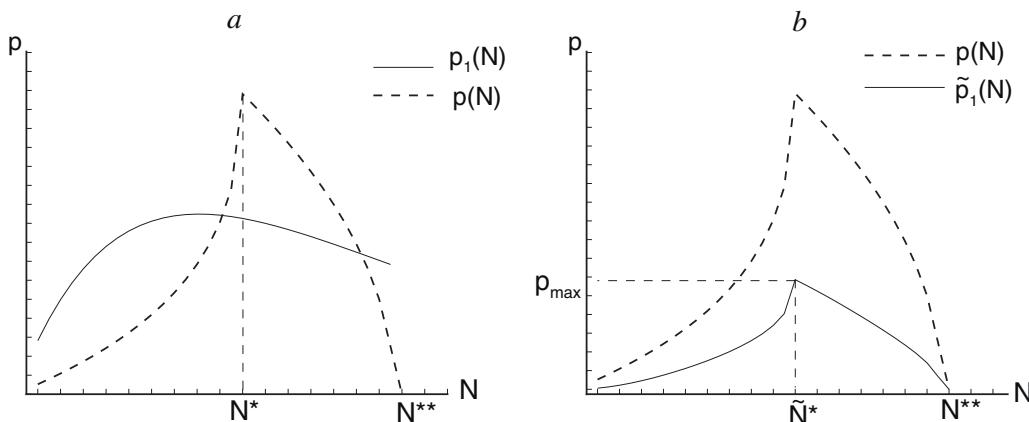


Fig. 11. (a) The probability $p_1(N)$ of having a single trap located within the peak sub-domain when the traps are randomly distributed over the unit interval. (b) The resulting probability $\tilde{p}_1(N)$ of having the error with the given range (11) when a single trap falls into the peak sub-domain. The probability $p(N)$ of having the error (10) with the given range (11) is shown as a dashed line in Fig. 10a and b.

distribution of traps. Here \tilde{p}_m , $m = 2, 3, \dots, N$, is the probability of having the error with the given range (11) when m traps fall into the peak sub-domain. We would then have to compute the total probability $\tilde{P}(N) = \sum_{m=1}^N \tilde{p}_m$ and see if $\tilde{P}(N) \approx 1$. Computation of the probability $\tilde{P}(N)$ is beyond the scope of this paper and will become a topic of our future work. However, let us note that in Section 2 we have shown that, in the case $\delta \ll 1$ (which is the main focus of this paper), the probability of the event that more than one trap fall into the peak area is significantly less than 50% and getting smaller with a decrease in δ . Thus, having restricted our analysis to the case of high population aggregation (i.e. $\delta \ll 1$), corrections to Eq. (37) are expected to be small.

Therefore, based on our present results we believe that installing traps at nodes of an equidistant Cartesian grid is a better option than installing them randomly, when a pest abundance is evaluated for a high aggregation density distribution. Apart from the issue of accuracy discussed above, an equidistant grid of traps is a simpler and cheaper option than generating a random grid. A random grid of traps eliminates the bias error, but, as we already discussed in Section 2, the bias problem does not exist when a highly aggregated density distribution (a single peak) is considered. Finally, another important argument in favour of a grid of equidistant traps, is that such a grid is better suited for a multi-patch distribution. If we have a multi-patch density function where all patches have approximately the same width (i.e. a collection of several peaks scattered over the monitored area), then installing a grid with the number of traps $N > N_{\text{threshold}}$ will detect all the patches (Petrovskaya and Embleton, 2013), while we cannot guarantee the same result for randomly distributed traps.

It should be mentioned in our discussion that a quadratic approximation of the density function introduces an interpolation error that will affect the reliability of our conclusions if a peak function different from quadratic is considered. In other words, the number $N_{\text{threshold}}$ for a function different from a quadratic polynomial, will differ from an estimate obtained for the quadratic function. However, numerical test cases studied in Petrovskaya and Embleton (2013) showed that an estimate of $N_{\text{threshold}}$ obtained for a quadratic function can in most cases be relied upon as the threshold number of traps, if another highly aggregated density distribution is considered. Therefore, in our opinion, the advantages of an equidistant grid of traps are generic and hence should be kept in mind when a decision about the trapping procedure is made.

7. Concluding remarks

We have considered a problem of pest insect population size evaluation from spatially discrete data in the case that the population is localised in a small sub-domain. In order to estimate the true mean density of the pest population, we used the statistical method (2). This method is popular among ecologists and is widely employed for obtaining estimates of the pest insect population; however, to the best of our knowledge, it is normally used under assumption that the spatial distribution of the monitored population is approximately uniform. Meanwhile, there are many ecological situations when species are highly aggregated in space. One example is given by the early stage of biological invasion (Shigesada and Kawasaki, 1997) but there are other mechanisms as well (Okubo, 1986).

The main results of our paper are as follows:

- It has been shown in the paper that a standard methodology does not work when the density of a highly aggregated pest population – a “peak distribution” – is measured by a trapping procedure where a small number N of traps is installed. Namely,

if the number of traps is small, an estimate of the mean density becomes a random variable with an error of high magnitude and we cannot control the accuracy of evaluation.

- We have obtained a probability of an accurate estimate based on the assumption that only one trap falls within a sub-domain where the pest population density is different from zero. The probability of an accurate estimate has been calculated for the one-dimensional and the two-dimensional problem.
- It has been revealed that, under the above assumption of a single trap within a peak sub-domain, there is a critical number N^* of traps that gives us the maximum probability of obtaining an accurate estimate of the mean density. A further increase in the number of traps may reduce the probability of an accurate answer.
- It has been shown that, for the same number of traps $N \sim 1/\delta$ (where δ is the width of the population peak), use of regular trap grids is more effective compared to random grids.

Note that our analysis remains valid and hence the findings apply to a somewhat more general case when the information about the local species density, i.e. at the location of the grid nodes, is obtained not through trapping but by using another sampling technique. The details of the technique may depend on the biological traits of the monitored species. For instance, while the density of flying and walking/crawling insects is indeed usually estimated from trap counts, the density of soil pest invertebrates is estimated by taking soil samples (Blackshaw, 1983). Whatever is the particular technique, the expected outcome of the sampling procedure is an estimate of the population density at the location of the samples/traps.

Our first finding from the above list demonstrates that the conventional conclusions that ecologists make from the use of the method (2) should be regarded with care, especially when high aggregation density distributions are considered. Ecologists are well aware that there is uncertainty in the estimation of the pest abundance, which may become worse as the number of samples decreases (Binns et al., 2000). However, the idea of handling an error in the mean density evaluation as a random variable has not been discussed in the ecological literature so far. Taking into account the risk factor related to the uncertainty in the mean density evaluation when the number N of traps/samples is small may result in the revision of an appropriate methodology for decision making, especially in the case that high aggregation population density is likely to occur.

An interesting question is to what extent our approach remains valid in case there are multiple patches of pests across the field, a situation which is often observed in reality (Barclay, 1992). In the present paper we have considered pest population density distributions such that the entire population is localised to a single patch within an agricultural field. Whilst this kind of distribution has ecological significance as it corresponds to the early stage of biological invasion, it is somewhat of an extreme case. Meanwhile, the technique we have presented in the paper to analyse the single patch distribution can also be used in the multi-patch case. Assuming the patches are on average the same size, it is then a matter of multiplying the probability p of achieving an accurate estimate for a single patch, by the total number of patches. When using a random distribution of traps, this problem is complex as there is no way of knowing how many patches there are in the field. However, if the traps are placed uniformly and the number of traps is sufficient to detect one patch, then it will be sufficient to detect them all. Of course it may be that the number of traps needed in this case is too large to be practical. Analysis should be performed to determine whether in this situation a smaller number of randomly distributed traps would in fact be more suitable. This will become a focus of future research.

Finally, another important direction of research is to take into account the uncertainty (noise) in the data obtained from a trapping procedure. A more detailed consideration of this issue should also become a focus of future work.

Appendix A. The probability $p(N)$ in the 2D case

Consider N traps installed over the domain, where we assume that only one trap is located within the peak sub-domain D_p , and any other traps fall in the region outside D_p where the density distribution is zero. The location of this single trap is denoted $\mathbf{r}_0 = (x_0, y_0)$, and is parameterised as

$$x_0 = r \cos \theta + x^*, \quad y_0 = r \sin \theta + y^*, \quad (38)$$

where $r \in [0, R]$ and $\theta \in [0, 2\pi]$. The location of \mathbf{r}_0 is randomised by considering r and θ as uniformly distributed random variables. The population density at this location, written as $u(x_0, y_0) \equiv u_0$, is then calculated as

$$u_0 \approx g(x_0, y_0) = B - A((x_0 - x^*)^2 + (y_0 - y^*)^2) = A(R^2 - r^2), \quad (39)$$

$r \in [0, R]$,

where we have used the fact that $B = AR^2$. The mean density $M(N)$ is then

$$M(N) = \frac{u_0}{N} = \frac{A(R^2 - r^2)}{N}.$$

The true mean density E is computed as

$$E = \int_0^1 \int_0^1 u(x, y) dx dy = \frac{1}{2} A \pi R^4.$$

Once again we require that the error (10) is sufficiently small, therefore we impose condition (11). From the above values of $M(N)$ and E we obtain

$$\frac{(1-\tau)A\pi R^4}{2} \leq \frac{A(R^2 - r^2)}{N} \leq \frac{(1+\tau)A\pi R^4}{2}. \quad (40)$$

Let us first consider the upper limit in (40), namely

$$\frac{A(R^2 - r^2)}{N} \leq \frac{(1+\tau)A\pi R^4}{2}. \quad (41)$$

Solving for r we obtain

$$r \geq r_I = R \sqrt{1 - \frac{N(1+\tau)\pi R^2}{2}}, \quad (42)$$

where r_I exists for

$$N \leq N_{2D}^* = \frac{2}{(1+\tau)\pi R^2}. \quad (43)$$

We now consider the inequality

$$\frac{(1-\tau)A\pi R^4}{2} \leq \frac{A(R^2 - r^2)}{N}. \quad (44)$$

After some rearrangement we arrive at

$$r \leq r_H = R \sqrt{1 - \frac{N(1-\tau)\pi R^2}{2}}. \quad (45)$$

The limit r_H exists when

$$N \leq N_{2D}^{**} = \frac{2}{(1-\tau)\pi R^2}. \quad (46)$$

As $\tau \in (0, 1)$, the number $N_{2D}^* < N_{2D}^{**}$. We consequently have three cases to consider when evaluating the probability $p(N)$ that the error (10) is within the prescribed tolerance τ .

Case 1: $N \leq N_{2D}^*$.

For this range of N , both r_I and r_H exist. Therefore the admissible range of the parameter r is

$$r_I \leq r \leq r_H. \quad (47)$$

Since r is a uniformly distributed random variable the probability $p(N)$ or $M(N)$ being sufficiently close to the true mean density E can be computed as

$$p_I(N) = \frac{r_H - r_I}{r_{\max} - r_{\min}} = \frac{r_H - r_I}{R},$$

where $r_{\min} = 0$ and $r_{\max} = R$. From (42) and (45) we thus have

$$p(N) = \sqrt{1 - \frac{N(1-\tau)\pi R^2}{2}} - \sqrt{1 - \frac{N(1+\tau)\pi R^2}{2}}. \quad (48)$$

Case 2: $N_{2D}^* < N \leq N_{2D}^{**}$.

In this instance, r_I no longer exists, but the inequality (41) always holds. Therefore the lower limit in (47) should be replaced by $r_{\min} = 0$. The admissible range now becomes $0 \leq r \leq r_H$, therefore the probability $p(N)$ is described by

$$p_{II}(N) = \frac{r_H - 0}{r_{\max} - r_{\min}} = \frac{r_H}{R}.$$

Substituting in the values for r_I and r_H we arrive at

$$p(N) = \sqrt{1 - \frac{N(1-\tau)\pi R^2}{2}}. \quad (49)$$

Case 3: $N > N_{2D}^{**}$.

When the number of traps N exceeds the limit N_{2D}^{**} , neither r_I nor r_H exist, and the inequalities (41) and (44) never hold. There is thus no admissible range of r for this range of N . The probability that the error (10) is sufficiently small is then $p_{III}(N) = 0$.

References

- Alexander, C.J., Holland, J.M., Winder, L., Wooley, C., Perry, J.N., 2005. Performance of sampling strategies in the presence of known spatial patterns. *Annals of Applied Biology* 146, 361–370.
- Barclay, H.J., 1992. Modelling the effects of population aggregation on the efficiency of insect pest control. *Researches on Population Ecology* 34, 131–141.
- Binns, M.R., Nyrop, J.P., Werf, W., 2000. Sampling and Monitoring in Crop Protection: The Theoretical Basis for Developing Practical Decision Guides. CABI, Wallingford.
- Blair, A., Zahm, S.H., 1991. Cancer among farmers. *Occupational Medicine – State of the Art Reviews* 6, 335–354.
- Blackshaw, R.P., 1983. The annual leatherjacket survey in Northern Ireland, 1965–1982, and some factors affecting populations. *Plant Pathology* 32, 345–349.
- Bliss, C.I., 1941. Statistical problems in estimating populations of Japanese beetle larvae. *Journal of Economic Entomology* 34, 221–232.
- Burn, A.J., 1987. Integrated Pest Management. Academic Press, New York.
- Byers, J.A., Anderbrant, O., Lofqvist, J., 1989. Effective attraction radius: a method for comparing species attractants and determining densities of flying insects. *Journal of Chemical Ecology* 15, 749–765.
- Chapman, J.W., Smith, A.D., Woiwod, I.P., Reynolds, D.R., Riley, J.R., 2002. Development of vertical-looking radar technology for monitoring insect migration. *Computers and Electronics in Agriculture* 35, 95–110.
- Cox, M.G., 2007. The area under a curve specified by measured values. *Metrologia* 44, 365–378.
- Davis, P.J., Rabinowitz, P., 1975. Methods of Numerical Integration. Academic Press, New York, USA.

- 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, FL, pp. 33–54.
- Dent, D., 1991. *Insect Pest Management*. CABI, Wallingford.
- Ester, A., Rozen, K., 2005. Monitoring and control of *Agriotes lineatus* and *A. obscurus* in arable crops in the Netherlands. *IOBC Bulletin 'Insect Pathogens and Insect Parasitic Nematodes'* 28, 81–86.
- Ferguson, A.W., Klukowski, Z., Walczak, B., Clark, S.J., Mugglestone, M.A., Perry, J.N., Williams, I.H., 2003. Spatial distribution of pest insects in oilseed rape: implications for integrated pest management. *Agriculture Ecosystems and Environment* 95, 509–521.
- Foster, S.P., Harris, M.O., 1997. Behavioral manipulation methods for insect pest management. *Annual Review of Entomology* 42, 123–146.
- Hagler, J.R., Jackson, C.G., 2001. Methods for marking insects: current techniques and future prospects. *Annual Review of Entomology* 46, 511–534.
- Holland, J.M., Perry, J.N., Winder, L., 1999. The within-field spatial and temporal distribution of arthropods in winter wheat. *Bulletin of Entomological Research* 89, 499–513.
- Jepson, P.C., Thacker, J.R.M., 1990. Analysis of the spatial component of pesticide side-effects on non-target invertebrate populations and its relevance to hazard analysis. *Functional Ecology* 4, 349–355.
- Karandinos, M.G., 1976. Optimum sample size and comments on some published formulae. *Bulletin of the Entomological Society of America* 22, 417–421.
- Legg, D.E., Moon, R.D., 1994. Bias and variability in statistical estimates. In: Pedigo, L.P., Buntin, G.D. (Eds.), *Handbook of Sampling Methods for Arthropods in Agriculture*. CRC Press, Boca Raton, FL, pp. 55–69.
- Lincoln, F.C., 1930. Calculating waterfowl abundance on the basis of banded returns. *USDA Circular* 118, 1–4.
- Malchow, H., Petrovskii, S.V., Venturino, E., 2008. *Spatiotemporal Patterns in Ecology and Epidemiology: Theory, Models, and Simulations*. Chapman & Hall/CRC Press, London.
- Mei, L., Guan, Z.G., Zhou, H.J., Lv, J.Z., Zhu, R., Cheng, J.A., Chen, F.J., Löfstedt, C., Svanberg, S., Somesfalean, G., 2012. Agricultural pest monitoring using fluorescence lidar techniques. Feasibility study. *Applied Physics B: Lasers and Optics* 106, 733–740.
- Metcalf, R.L., 1982. *Introduction to Insect Pest Management*. Wiley, London.
- Murray, J.D., 1989. *Mathematical Biology*. Springer, Berlin.
- Okubo, A., 1986. Dynamical aspects of animal grouping: swarms schools flocks and herds. *Advances in Biophysics* 22, 1–94.
- Perry, J.N., 1989. Review: population variation in Entomology: 1935–1950. I. Sampling. *The Entomologist* 108, 184–198.
- Petrovskaya, N.B., Petrovskii, S.V., 2010. The coarse mesh problem in ecological monitoring. *Proceedings of the Royal Society A* 466, 2933–2953.
- Petrovskaya, N.B., Venturino, E., 2011. Numerical integration of sparsely sampled data. *Simulation Modelling Practice and Theory* 19, 1860–1872.
- Petrovskaya, N.B., Petrovskii, S.V., Murchie, A.K., 2012. Challenges of ecological monitoring: estimating population abundance from sparse trap counts. *Journal of the Royal Society, Interface* 9, 420–435.
- Petrovskaya, N.B., Embleton, N.L., 2013. Evaluation of peak functions on ultra-coarse grids. *Proceedings of the Royal Society A*, accepted for publication.
- Petrovskii, S.V., Morozov, A.Y., Venturino, E., 2002. Allee effect makes possible patchy invasion in a predator-prey system. *Ecology Letters* 5, 345–352.
- Petrovskii, S.V., Li, B.-L., Malchow, H., 2003. Quantification of the spatial aspect of chaotic dynamics in biological and chemical systems. *Bulletin of Mathematical Biology* 65, 425–446.
- Petrovskii, S.V., Malchow, H., Hilker, F.M., Venturino, E., 2005. Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control. *Biological Invasions* 7, 771–793.
- Petrovskii, S., Bearup, D., Ahmed, D.A., Blackshaw, R., 2012. Estimating insect population density from trap counts. *Ecological Complexity* 10, 69–82.
- Raworth, D.A., Choi, W.-J., 2001. Determining numbers of active carabid beetles per unit area from pitfall-trap data. *Entomologia Experimentalis et Applicata* 98, 95–108.
- Reisen, W.K., Lothrop, H.D., 1999. Effects of sampling design on the estimation of adult mosquito abundance. *Journal of the American Mosquito Control Association* 15, 105–114.
- Seber, G.A., 1982. *The Estimation of Animal Abundance and Related Parameters*. Charles Griffin, London, UK.
- Snedecor, G.W., Cochran, W.G., 1980. *Statistical Methods*. The Iowa State University Press, Ames.
- Silver, J.B., 2008. *Mosquito Ecology: Field Sampling Methods*. Springer, NY.
- Shigesada, N., Kawasaki, K., 1997. *Biological Invasions: Theory and Practice*. Oxford University Press, Oxford.
- Southwood, T.R.E., Henderson, P.A., 2000. *Ecological Methods*. Blackwell Science.
- Stern, V.M., 1973. Economic thresholds. *Annual Review of Entomology* 18, 259–280.
- Spiegel, M.R., 1992. *Theory and Problems of Probability and Statistics*. McGraw-Hill, New York, pp. 108–109.
- Turchin, P., 2003. *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton University Press, Princeton.
- Vlug, H.J., Paul, H., 1986. Sampling leatherjackets. *Mededelingen van de Faculteit Landbouwwetenschappen, Rijksuniversiteit Gent* 51, 939–942.
- Ward, S.A., Rabbinge, R., Mantel, W.P., 1985. The use of incidence counts for estimation of aphid populations. I. Minimum sample size for required accuracy. *European Journal of Plant Pathology* 91, 93–99.