Entomology in Focus 1

Cláudia P. Ferreira
Wesley A.C. Godoy *Editors*

Ecological Modelling Applied to Entomology

Chapter 8
Computational Methods for Accurate Evaluation of Pest Insect Population Size

Natalia Petrovskaya and Nina Embleton

Abstract Ecological monitoring aims to provide estimates of pest insect abundance, where the information obtained as a result of monitoring is then used for making decisions about means of control. In our paper we discuss the basic mathematics behind evaluating the pest insect abundance when a trapping procedure is used to collect information about pest insect species in an agricultural field. It will be shown that a standard approach based on calculating the arithmetic average of local densities is often not the most efficient method of pest population size evaluation and more accurate alternatives, known as methods of numerical integration, can be applied in the problem. A mathematical background for methods of numerical integration on regular grids of traps will be provided and examples of their implementation in ecological problems will be demonstrated. We then focus our attention on the issue of pest abundance evaluation accuracy when data available in the problem are sparse and consider the extreme case when the uncertainty of evaluation is so big, that an estimate becomes a random value. We complete our discussion with the consideration of irregular grids of traps where numerical integration techniques can also be applied.

Keywords Ecological monitoring • Numerical integration • Pest control • Coarse grid

8.1 Introduction

Pests are a sustained and significant problem in the production of food across the globe. The term ‘pest’ can be used to describe any organism which is deemed to cause harm to mankind in some manner; in crop production this label is given to those which damage or destroy potential produce to an unacceptable extent. In many

N. Petrovskaya (✉) • N. Embleton
School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK
e-mail: n.b.petrovskaya@bham.ac.uk; embleton@for.mat.bham.ac.uk

C.P. Ferreira, W.A.C. Godoy (eds.), *Ecological Modelling Applied to Entomology*, Entomology in Focus 1, DOI 10.1007/978-3-319-06877-0_8

ecologically important cases the definition above implies that an agglomeration of organisms is considered: for example, while one or two occasional insects do not make any significant harm to the crop in an agricultural field, the damage to the crop can become dramatic if the number of insects exceeds a certain threshold. Hence, in many ecological situations the definition of the term pest also requires the definition of the *pest abundance* or the *pest density distribution* in the spatial domain under consideration.

Crops are vulnerable to attack from pests both during the growing process and after they have been harvested. When pests of crops prior to harvest are considered, the focus is often predominantly on arthropods, plant pathogens and weeds (e.g. Louws et al. 2010; Ruberson 1999). Estimates of the annual worldwide loss due to pests at this stage in the production process lie between 35 and 42 % (Oerke 2006; Pimentel 1997). In particular, the pre-harvest loss of 14–15 % of the world's crops has been attributed to insect pests (Pimentel 2009; Pimentel and Pimentel 2008). Further losses are incurred after the crops have been harvested. This can be due to infestation of stored crops by pests such as insects, rodents, birds, as well as micro-organisms which cause damage both quantitative and qualitative in nature (Gwinner et al. 1996). Such losses have been estimated to range from 10 to 25 % (Pimentel and Pimentel 2008).

8.1.1 Basic Principles of Integrated Pest Management

Pest management has the obvious goal of preventing or minimising the damage pests cause to crops and various approaches have been used to achieve this goal. Measures of so-called 'preventative pest management' can be put into practise; the idea being to try to stop the pest population from becoming a problem in the first place. Age-old examples of such a tactics are crop rotation and intercropping. In a crop rotation, instead of an agricultural field consistently being used to grow the same crop, different crops which critically host different pests, are grown sequentially. Intercropping is the planting of different crops within the same field at the same time. Variety can also be introduced by planting several genotypes of the same crop species within a field. Introducing heterogeneity in such ways, either spatially, temporally, or genotypically, can destabilise the life cycle of a pest and has been documented to help to control pest populations (Liebman and Dyck 1993; Shoffner and Tooker 2013). A pest's preference for a certain plant can be exploited to the farmer's advantage using a technique called trap cropping. Here, crops are interspersed with plants that are more attractive to the pest and thus act as sacrificial decoys. This diversionary ploy can be sufficient to protect the crop in itself, otherwise, it reduces the area of the field to be subjected to further management tactics should they be needed since the pests are then located in some field sub-domain (Hokkanen 1991). Another precautionary measure is to grow crops which have been cultivated to be resistant to pest attack. Grafting has been used for centuries to manage certain pathogens and it has also been deemed to be useful in

the control of arthropod pests and weeds (Louws et al. 2010). A more scientifically advanced means of pest resistant plant cultivation is genetic modification. This is a relatively recent initiative of which the risks are not yet fully understood, however, its potential to become the dominant pest management strategy has certainly been recognised and consequently it has become the focus of much research (e.g. Bates et al. 2005; Christou et al. 2006; Gatehouse et al. 2011; Smigocki et al. 2013).

Another way of managing pests is to implement a control action, that is, to employ a means of killing the pest organisms. The most widely used control action is the application of pesticides. It has been estimated that around 3×10^9 kg are used across the globe per year (Pimentel 2009). Biological control actions, e.g. releasing a natural enemy of the targeted pest into the agroecosystem, provide an alternative to the use of chemicals. However, the indiscriminate use of control actions or using them as a preventative measure can have serious negative consequences. For instance, the regular use of pesticides often leads to the pest becoming resistant making future management a more difficult task (Alyokhin et al. 2008). Another unwanted side effect can be that the pesticide has lethal or sub-lethal effects on natural enemies (Sohrabi et al. 2013) which can cause a resurgence in the pest population or a secondary pest to emerge.

Recognition that precautionary tactics are rarely sufficient to manage pests alone and that relying entirely on control action is not a durable approach led to the emergence of the concept of *integrated pest management* (IPM) (Kogan 1998). IPM is the incorporation of several different tactics which work cooperatively together to protect crops from pest attack in a more sustainable way. It consists of the three phases. Firstly, preventative measures of pest management are put into place. Subsequently, the pest abundance is monitored. The decision of whether or not to implement a control action is then made by comparing the abundance of pests against some threshold level, i.e. the limit at which intervening becomes worth the effort or expense. Such threshold values can be decided upon by taking a variety of factors into consideration. The most often used are economic thresholds (Stern et al. 1959) as usually the overriding concern is that the pest management programme is financially viable (e.g. see Higley and Pedigo 1996). The basic principle of IPM is therefore that a control action is only used if and when it is necessary. Thus monitoring is key to the decision process and is considered an essential part of any integrated pest control and management programme (Burn et al. 1987; Metcalf and Luckmann 1982).

8.1.2 Monitoring Methodologies in IPM

A correct choice of a monitoring methodology is very important for the success of an IPM programme. Since different pest types have different behaviours, the monitoring methodologies in IPM programmes vary accordingly. We thus limit our scope to the consideration of insect pests; henceforth in the text the generic term ‘pest’ is used synonymously with ‘insect pest’ unless otherwise stated. The

procedure also depends on the environment to be monitored. In our paper we consider pest management of crops prior to harvest, where we take the spatial scale of the monitoring procedure to be that of an agricultural field. A complete census in this case is hardly practical or indeed possible, therefore the population abundance must instead be estimated. The data to form such an estimate is collected by sampling the pest population for which there exists a multitude of techniques (Ausden 1996; Blackshaw 1983; Hutchins 1994; Mayor and Davies 1976; Southwood and Henderson 2000).

A direct, in-situ count can be made of the number of pests in a sample unit e.g. a plant or a unit area of habitat. For the more inconspicuous species, the counting process can be made easier by dislodging the pests from the plant using a practise known as 'knockdown'. In some instances a sample of the habitat itself may be carefully removed and taken to a laboratory where the count can then be made.

Once the data has been collected the arithmetic mean number of pests M per sample unit is calculated as follows:

$$M = \frac{1}{K} \sum_{k=1}^{K} f_k, \tag{8.1}$$

where f_k are the individual sample counts, and K is the number of sample counts taken (Davis 1994). From the mean number of pests per unit area, an estimate of the number of pests in the entire agricultural field is obtained by multiplying by the area of the field (Snedecor and Cochran 1980). A mean number of pests per plant can be converted to the mean per unit area by multiplying by the mean number of plants in such an area. Such an estimate of pest abundance is considered an 'absolute' estimate since the sample counts directly reflect the number of pests in the sample unit.

Alternatively samples can be taken via netting. For example, a net can be swung into the crops for a prescribed time or number of sweeps. The number of pest insects caught inside is then counted (e.g. see Pedigo and Rice 2009; Southwood and Henderson 2000). Netting is often used to sample insects on large agricultural fields, because it is quicker and more cost effective than inspection of individual plants.

Another widely used sampling technique is trapping. Traps are installed in the field, exposed for a certain amount of time (e.g. for a week), after which the traps are emptied and the pests counted. The position of the traps can be arbitrary; some ecologists opt for random grids of traps or choose appropriate sampling patterns (Alexander et al. 2005; Mayor and Davies 1976), but in many cases they are placed at the nodes of a rectangular grid (Ferguson et al. 2000; Holland et al. 1999). The traps can either be active, whereby an attractant is used to draw the pests into the traps e.g. bait or a pheromone, or they can be passive where capture relies on the activity of the pest species. The trap counts provide information about the pest population density at the position of the traps (Byers et al. 1989; Raworth and Choi 2001) and the sample mean density can then be calculated by scaling (8.1)

with relation to the area of the agricultural field, where f_k are now the pest densities at the sample locations.

The above techniques yield a relative estimate of the mean pest density rather than an absolute estimate. The counts are not a direct measure of pest abundance but are relative to the efficiency of the netting or trapping technique and the conditions at the time of sampling. Therefore only relative estimates which have been obtained via the same sampling technique and in the same conditions can be compared. It is possible to convert an estimate that is relative to an absolute estimate using regression analysis (Browde et al. 1992) or through calibration using experimental data (Evans et al. 1983). Steps to achieve this via mathematical modelling have also been made (Petrovskii et al. 2012).

An estimate of the population abundance can also be achieved using mark-release-recapture methods. Initial sampling is performed and the catch is counted and marked in some way (Hagler and Jackson 2001). The marked population is then released back into the agroecosystem and another round of sampling is conducted. An estimate of the population size can then be formulated using the condition that the proportion of marked insects in the field is equal to the proportion of marked insects found in the second sample. That is, the following can be rearranged to solve for I

$$\frac{I_M}{I} = \frac{I_{\tilde{M}}}{\tilde{I}}, \tag{8.2}$$

where I_M is the total number of marked insects, I is the number of insects in the entire population, $\tilde{I}$ is the number of insects caught in the second sample and $I_{\tilde{M}}$ is the number of those which are marked. This method works well in scientific studies but can hardly be afforded in nation-wide monitoring programmes as it requires considerable additional effort (such as insect marking and recapture).

8.1.3 The Problem of Accurate Estimation of Pest Abundance

Once an estimate of the pest population size or the mean pest density in an agricultural field has been acquired, a decision is made by comparing it to some threshold value(s). Let us consider the simplest case where a single threshold value is used. If the estimate falls below the threshold the decision is to take no action, whereas if it exceeds the threshold the decision is to intervene and implement a control action (e.g. see Binns et al. 2000, Chapter 1). Such action can, for instance, be the application of pesticides (Ester and van Rozen 2005; Stern 1973). Clearly the accuracy of the estimate is important in ensuring the correct decision is made, with the accuracy becoming particularly vital when it is close to the threshold value. An underestimate could mean action is not taken when it is needed leading to the loss of crops. Even with the use of pesticides the value of crops lost in the field to

pests has been estimated to be $2,000 billion per year (Pimentel 2009).[1] Obtaining a more accurate estimate of the pest abundance could lead to the more timely use of a control action and ultimately reduce crop loss.

On the other hand an overestimate could lead to a control action being used unnecessarily. Application of pesticides is costly and brings considerable damage to the environment (Jepson and Thacker 1990). Pesticides are known to contribute to air, soil and water pollution whilst there is growing evidence linking their use to human illnesses (Alavanja et al. 2013; Pimentel and Greiner 1997). It has been estimated that less than 0.1 % of pesticides used reach their targeted pest, the remaining 99.9 % is absorbed by some means into the environment (Pimentel 1995). Some of the loss occurs during application with the spray drifting outside of the intended area, however once applied to a crop, pesticides can then vaporise into the air, end up in surface or groundwater, be absorbed by plants or ingested by non-target species, or indeed remain in the soil. Furthermore, unnecessary application of pesticides is undesirable from an economic perspective; around $40 billion is spent per year applying pesticides (Pimentel 2009).

It is obvious from the above that there is a significant need for reliable methods to accurately evaluate the pest population size in order to avoid making an unjustified decision about control action. It is worth noting here that the accuracy required by pest monitoring is not always very demanding as it differs according to the purpose. In routine monitoring an error range can be 20–100 % (Pascual and Kareiva 1996; Sherratt and Smith 2008), whereas monitoring for research purposes can demand a higher degree of accuracy of 10 % (e.g. see Pedigo and Rice 2009, p. 245).

Several means of optimising the accuracy of an estimate have been considered in the ecological literature. One way is to ensure that the size of the data set is large enough i.e. that enough sample units are taken. It follows from Eq. (8.1) that the exact value of the population size will be obtained for infinitely large number K. Hence we can expect better accuracy of the estimate when K gets larger. A pre-sample (or series of them) can be used to obtain a sample mean and sample variance from which an estimate of the number of sample units needed to achieve a specified precision can be calculated (e.g. see Binns et al. 2000; Dent 2000; Pedigo and Rice 2009). However there is a trade-off between the number of sample units needed to achieve sufficient accuracy and the number that can be practically afforded. For instance, if a trapping procedure is applied in ecological research, the number K of traps per given area can be made quite large, e.g. in the order of hundreds. Meanwhile in routine pest monitoring programmes K rarely exceeds 20 (Mayor and Davies 1976) per a typical agricultural field with a linear size of several hundred meters and, in some cases, it can be as small as one or a few traps per field (Northing 2009). There are several practical reasons why the number of traps cannot be made large. An increase in the number of sample units equates to an increase in the amount of labour and hence finances required. In any real-world scenario

[1]The work Pimentel (2009) refers to pests in the generic sense of the term, i.e. insects, plant pathogens and weeds.

there is a limit to such resources. Also, traps introduce a disturbance into the field and installing a large number of them can damage the corresponding agricultural product. Furthermore, trapping imposes a disturbance on the pests which can in turn affect the results of the trapping technique, therefore from this perspective the number of traps should be minimised.

The efficacy of a sampling technique is also important to the accuracy of an estimate of the pest abundance. Means of sampling a pest population are constantly being reviewed leading to sampling equipment being developed and improved (Birmingham et al. 2011; Taboada et al. 2012). Another key consideration is the sampling plan, that is, the prescribed locations at which samples are to be taken. For an estimate to be accurate the sample must capture sufficient information to adequately represent the true pest presence. If conditions are homogeneous across the field, insects can be randomly distributed, however they often exhibit an aggregated spatial distribution (Ferguson et al. 2000; Holland et al. 1999). The sampling plan thus becomes crucial; it is important to avoid bias stemming from samples being placed entirely in areas where the pests are clustered, or likewise, entirely in areas of zero density. Comparisons of various patterns e.g. random, transects, quadrats, etc. have been made in order to make recommendations (Alexander et al. 2005).

8.1.4 Goals and the Road Map

As the accuracy of evaluation of the pest abundance remains a crucial issue in IPM programmes, any new method that can increase the accuracy must be carefully studied and its advantages and disadvantages must be documented in order to decide whether or not the method can be used in routine monitoring. Although ensuring a sufficiently accurate estimate has been considered in the ecological literature as discussed above, to our best knowledge the focus has predominantly been on how the data is collected. In our paper we instead look at the way in which the data is processed and discuss numerical integration techniques that present an alternative approach to the existing statistical methods. In recent years intensive study of numerical integration methods for ecological applications has been carried out (Embleton and Petrovskaya 2013, 2014; Petrovskaya and Embleton 2013; Petrovskaya et al. 2013; Petrovskaya and Petrovskii 2010; Petrovskaya et al. 2012) and in this book chapter we summarize our experience with the application of numerical integration methods to ecological problems. We will focus our attention on a trapping procedure made in a single agricultural field and on the evaluation of the total pest population size from the information provided by trap counts, but the results of our discussion can be extended to other sampling techniques.

The main goal of the book chapter is twofold. Firstly, we would like to draw the attention of our readers to methods of numerical integration as a reliable alternative to a standard statistical method (8.1). We therefore explain a mathematical background for numerical integration techniques, elaborate on how to apply

them in ecological problems and demonstrate that advanced numerical integration methods can often be more effective in the evaluation of pest abundance than the method (8.1).

Secondly, we want to discuss the issues of accuracy for various methods of numerical integration and to identify the main factors that may affect the accuracy. It will be shown that the accuracy of numerical integration depends on the number of traps available in the problem and we therefore often have to deal with a numerical integration problem where the data are sparse (see Sect. 8.1.3). Meanwhile, if the number of traps is fixed in the problem, a spatial pattern of the pest density distribution remains the most crucial factor that affects the accuracy of numerical integration and this is another key topic that we discuss in this book chapter.

While most of our study with regard to the issues above will be done for regular grids of traps, we are also interested in the study of accuracy on quasi-irregular and random grids, as sampling patterns that result in such grids are backed by ecologists as mentioned in Sect. 8.1.3. It is worth noting that, although the spatial pattern of the sample units is considered important when collecting the data, an estimate of pest abundance based on the sample mean does not use this information directly. It can readily be seen that the expression for the sample mean (8.1) has no spatial dependence. Alternatively, an estimate formulated by means of numerical integration uses the spatial distribution information and we will see the implications of this approach.

The chapter is organised as follows. In the next section, we briefly explain basic information about the theory of numerical integration. In Sect. 8.3 we introduce a coarse grid problem that may hamper the use of numerical integration methods in ecological applications. In Sects. 8.2 and 8.3 we consider standard examples that have no ecological meaning but serve the purpose of illustrating numerical integration techniques well. We then demonstrate in Sect. 8.4 how to use methods of numerical integration in order to evaluate the total population size from discrete spatial data on regular grids. We also check the accuracy of various numerical integration methods by applying them to spatial population distributions of different complexity and conclude that knowledge of a spatial pattern is the most important requirement when accuracy of numerical integration is concerned. In Sect. 8.5, we discuss highly aggregated density distributions that present the most difficult case for numerical integration methods. In Sect. 8.6, we investigate the effect of a grid's irregularity on the population size estimation. Finally, in section "Concluding Remarks" we summarise our experience with the numerical integration problem in pest insect monitoring and control.

8.2 Theory of Numerical Integration

In this section we provide a brief discussion of methods of numerical integration and explain basic concepts related to this technique. We introduce a generic problem of numerical integration and elaborate on the accuracy of integration when various

methods are considered. For the sake of simplicity our discussion will mainly be focused on the one-dimensional case, but it can be readily extended to multi-dimensional problems.

8.2.1 Basic Concepts of Numerical Integration

Methods of numerical integration have to be applied when an integrand $f(x)$ defined over the interval $[a, b]$ is only given to us at a discrete set of points. This is a common situation when we make experimental measurements of the function $f(x)$ or when $f(x)$ is obtained as a result of computer simulation. If we consider the points $x_i, i = 1, \ldots, N+1$ where the function values $f_i \equiv f(x_i)$ are available, then computation of the integral

$$I = \int_a^b f(x)dx, \tag{8.3}$$

is reduced to computation of a weighted sum of the values f_i,

$$I \approx \tilde{I}(N) = \sum_{i=1}^{N+1} \omega_i f_i. \tag{8.4}$$

The basic problem of numerical integration is therefore to find weight coefficients ω_i such that the sum $\tilde{I}(N)$ will approximate the integral I with appropriate accuracy. The theory of numerical integration states that the weights ω_i in (8.4) depend on the number $N+1$ of points x_i where the function values f_i are available. Thus the accuracy requirement can be formulated for any numerical integration problem as

$$\tilde{I}(N) \to I, \text{as} N \to \infty, \tag{8.5}$$

and every time that the integration weights ω_i are defined in a new method of numerical integration, the condition (8.5) must be verified.

The condition (8.5) tells us that the weighted sum of function values (8.4) gets closer to the precise integral I when the number of points we use for integration increases. However, in order to come up with an efficient method of numerical integration we also want to know how fast the approximation $\tilde{I}(N)$ will approach the precise integral I when we increase N in (8.5). We have to introduce the concept of integration error in order to answer this question. Let us assume that the value I of the integral (8.3) is known to us. The integration error $E(N)$ is then defined as

$$E(N) = |I - \tilde{I}(N)|. \tag{8.6}$$

In many cases it is also convenient to consider the relative integration error,

$$e(N) = \frac{|I - \tilde{I}(N)|}{|I|}. \tag{8.7}$$

Consider now the unit interval $[0, 1]$. Let $x_1 = 0$ and equidistant points x_i, $i = 1, \ldots, N + 1$ be located over the interval, so that the distance between any two neighbouring points x_i and x_{i+1} is $h = 1/N = const < 1$. We refer to the set of points x_i, $i = 1, \ldots, N + 1$ as a *regular (or uniform) computational grid* of points. The points x_i are often called *grid nodes* and the distance h is referred to as *the grid step size*.

Once a computational grid has been generated, the integration error can be rewritten in terms of the distance h between neighbouring points as $E = E(h)$. We have $h \to 0$ as $N \to \infty$ and the condition (8.5) becomes

$$E(h) \to 0, \quad \text{as} \quad h \to 0, \tag{8.8}$$

where $E(h)$ is given by (8.6) after substituting $N = 1/h$. The formula (8.6) gives us a rigorous definition of the integration error, but it still remains unclear from (8.6) how we can check and control the condition (8.8), as the integral I is, of course, not available in real-life computations. Thus, instead of computing the exact value of $E(h)$ based on the exact value of the integral I, we make *an estimate* of the integration error in order to be able to check the condition (8.8). In the theory of numerical integration an integration error estimate is often considered in the following form Davis and Rabinowitz (1975)

$$E(h) = Ch^p, \tag{8.9}$$

where the constant C and the power p depend on a specific method of numerical integration used in the problem. The representation (8.9) of the integration error allows us to conclude about *the convergence rate*, i.e. to conclude how fast the error will decrease if we increase N. In other words, the formula (8.9) gives us the information on how fast $E(h) \to 0$, as $h \to 0$, and it is very important for our further discussion to emphasise here that h in the expression (8.9) is assumed to be small.

Let the integral be evaluated on a regular grid of $N_0 + 1$ points. The expression (8.9) reads that if we increase the original number N_0 as $N_1 = 2N_0$ then $h_1 = 1/N_1 = (1/2)h_0$ will be two times smaller, and the new error $E(h_1)$ will be 2^p times smaller. Obviously, the relative error (8.7) will exhibit similar behaviour. It is also obvious that the discussion above is true for any interval $[a, b]$ where the integrand $f(x)$ is considered. Indeed, any interval $[a, b]$ can be mapped onto the unit interval $[0, 1]$ by a linear transformation $x = (\hat{x} - a)/(b - a)$, where $x \in [0, 1]$ and $\hat{x} \in [a, b]$.

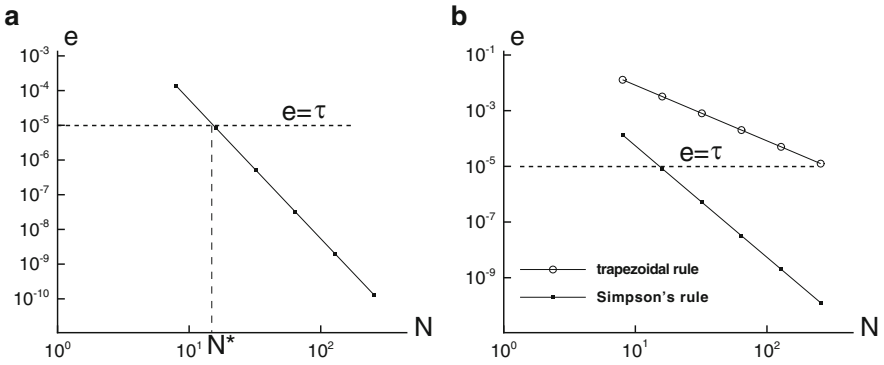

Fig. 8.1 The results of numerical integration of the function (8.10). (**a**) The convergence curve for the Simpson method of numerical integration. The relative integration error $e = e(N)$ is shown on the logarithmic scale. (**b**) Comparison of the convergence rate for the trapezoidal rule (*solid line*, *open circle*) and the Simpson rule (*solid line*, *closed square*)

From a practical viewpoint the concept of convergence means that we can control the accuracy of integral evaluation. This statement is illustrated in Fig. 8.1a, where the integral

$$I = \int_0^{\pi} \sin\, x dx \tag{8.10}$$

is evaluated by a selected method of numerical integration (composite Simpson's rule.[2]) Again, we assume that the function values are only available on a regular grid of $N + 1$ equidistant points. We start from the fixed number $N = 8$, and compute the approximate integral $\tilde{I} = \tilde{I}(N)$. As the precise integral $I = 2$ is known to us, we can compute the relative integration error (8.7). We then double the number N and repeat our computation of the error $e(N)$. After making this computation several times we obtain the error (8.7) as a function of N.

The graph $e(N)$ is shown in Fig. 8.1a on the logarithmic scale. It can be seen from the slope of the graph that the error decreases as h^4, where $h = 1/N$. The graph also gives us information about the threshold number N^* for which the following condition holds

$$e \leq \tau, \tag{8.11}$$

[2]The detailed description of this method is not important for our present discussion and will be provided later in the text.

where τ is a prescribed tolerance. If, for example, we choose $\tau = 10^{-5}$, then the accuracy (8.11) will be achieved for any $N \geq N^*$, where $N^* = 32$ (see Fig. 8.1a). Better accuracy requires a bigger number of points where function values are available, while larger τ (e.g., $\tau = 10^{-3}$) means that we can use a smaller number of points to evaluate the integral.

The convergence rate (8.9) of a particular method of numerical integration depends on the definition of the weight coefficients ω_i in the formula (8.4), and two different methods may therefore have different convergence rates. One example illustrating this statement is shown in Fig. 8.1b. In the figure we repeat the procedure previously explained for the graph in Fig. 8.1a, when another method of numerical integration (composite trapezoidal rule) is applied in the same problem. While we do not discuss here the definition of weight coefficients ω_i in each method, it can be readily concluded from Fig. 8.1b that the convergence rate of the trapezoidal rule is much slower than the convergence of the Simpson rule. The error in the composite trapezoidal rule decreases as h^2, while for the composite Simpson rule it decreases as h^4. Hence a much bigger number of points is required to achieve the accuracy $\tau = 10^{-5}$, if the composite trapezoidal rule is employed in the same problem of numerical evaluation of the integral (8.10).

The above discussion leads us to the conclusion that if we have several methods of numerical integration then the method that has the fastest convergence rate (8.9) must be employed in the problem and all other methods should be dismissed. Unfortunately, things are not so straightforward. Firstly, a fast convergence rate always comes at the price of the method's complexity, and methods that converge faster usually have more restrictions upon their implementation than the methods that converge slowly. The Simpson rule in the example above has a faster convergence rate, but it cannot be applied for an arbitrary number N and we should instead require that N is an even number in order to define weight coefficients for the Simpson rule. On the other hand, the trapezoidal rule has a slower convergence rate but it is more flexible and can be applied for arbitrary N. In practical applications the restrictions upon implementation of a specific method of numerical integration must be taken into account and that often results in the choice of a slower convergent method in the problem. Secondly, when we choose a method of numerical integration for the problem we solve, we need to understand how laborious the method is. In other words, it may happen that the desired accuracy will be achieved for a smaller number N but at the price of a very big number of computations, especially in multi-dimensional problems. In the latter case we should ask ourselves if we can come up with an alternative method of integral evaluation that may have a slower convergence rate but is easier to implement. Finally, and this is the most serious and difficult problem in numerical integration, the formula (8.9) may become invalid when the number $N + 1$ of points we have at our disposal is small. Other criteria should then be employed to compare two methods of numerical integration.

The above issues will be further discussed in the following sections. Their understanding will require us to give an explicit definition of the weight coefficients in numerical integration formulas. Below we consider the computation of weights ω_i in the integral approximation (8.4).

8.2.2 *Definition of Weight Coefficients in Various Methods of Numerical Integration*

Consider a regular computational grid of N sub-intervals in the domain $[a, b]$, i.e. consider points $x_1 = a$, $x_{i+1} = x_i + h$, $h = (b - a)/N$. As in the previous section we assume that the function values $f_i = f(x_i)$ are available at points x_i, $i = 1, \ldots, N + 1$. Numerical integration on regular grids with $h = const$ can be done by the application of well-known methods from the Newton-Cotes family of integration rules, the trapezoidal rule and the Simpson rule being, perhaps, the most famous. In this subsection we briefly review several methods of numerical integration that stem from various choices of weight coefficients ω_i in a generic formula (8.4) when regular grids are considered.

The problem of numerical integration is often thought of as a problem of finding the area under the curve $f(x)$. Thus the most straightforward and intuitively clear method is to take the function values $f_i \equiv f(x_i)$ at equidistantly spaced points x_i and to construct a rectangle with the sides $h = x_{i+1} - x_i$ and f_i. The area

$$a_i = hf_i, \tag{8.12}$$

gives us an approximation of the integral $I_i = \int_{x_i}^{x_{i+1}} f(x)dx$. Once the area a_i has been computed for each $i = 1, 2, \ldots, N$, the sum $S = \sum_{i=1}^{N} a_i$ is considered as an approximation of the integral $I = \int_a^b f(x)dx$. Such consideration is based on precise definition of a definite integral (8.3) as the limit of Riemann sums and the proof exists that the sum S will converge to the integral I as $N \to \infty$ (e.g. see Apostol 1974). It immediately follows from (8.4) and (8.12) that the weights are given by $\omega_i = h$ for any $i = 1, 2, \ldots, N$.

The approximation (8.12) is shown in Fig. 8.2a, where the function $f(x)$ is replaced by a constant $f_i \equiv f(x_i)$ on each subinterval $[x_i, x_{i+1}]$. It is clear that the approximation of the function by a constant is not very accurate, and we can improve it if we consider a straight line connecting points x_i and x_{i+1} (see Fig. 8.2b). The area a_i of each sub-interval is now given by

$$a_i = \frac{1}{2}h(f_i + f_{i+1}), \tag{8.13}$$

and again we compute the approximation to the integral as

$$I \approx S = \sum_{i=1}^{N} a_i. \tag{8.14}$$

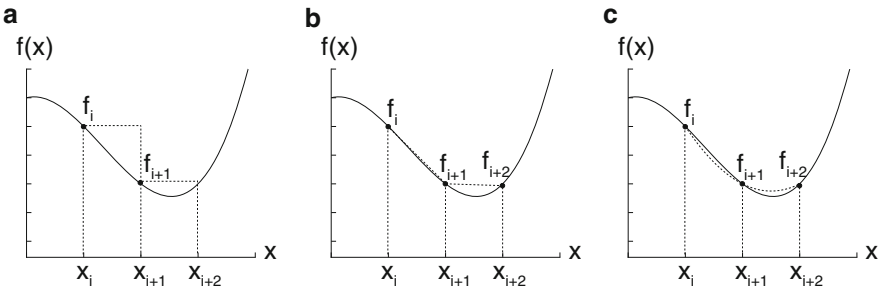

Fig. 8.2 Approximation of the function $f(x)$ by a polynomial of degree k. (**a**) Approximation by a constant ($k = 0$) over a subinterval $[x_i, x_{i+1}]$, (**b**) approximation by a straight line ($k = 1$), (**c**) three points x_i, x_{i+1} and x_{i+2} are required to approximate the function by a parabola ($k = 2$)

Substituting (8.13) into the sum (8.14) and re-arranging the terms, we arrive at the composite trapezoidal rule of integration,

$$I \approx S = \frac{h}{2}\left[f_1 + 2\sum_{i=2}^{N} f_i + f_{N+1}\right]. \tag{8.15}$$

The weight coefficients are now given by $\omega_1 = \omega_{N+1} = h/2$ and $\omega_i = h$, $i = 2, \ldots, N$.

Approximation of a function $f(x)$ by a straight line can be considered as replacing $f(x)$ by a linear polynomial on each subinterval $[x_i, x_{i+1}]$. If we go on with the idea of approximating the function by a polynomial of degree k, where $k = 0, 1, 2, 3, \ldots$, then our next step will be to consider $k = 2$ and to replace the integrand $f(x)$ by a quadratic polynomial. From a geometric viewpoint, this means drawing a parabola through three consecutive points where the function is defined. Clearly, we can use points x_i, x_{i+1}, x_{i+2} to define our quadratic polynomial as shown in Fig. 8.2c. The area under the curve is now approximated as the area a_i under the parabola passing through x_i, x_{i+1}, x_{i+2} and it is computed as

$$I_i = \int_{x_i}^{x_{i+2}} f(x)dx \approx \; a_i = \frac{1}{3}h(f_i + 4f_{i+1} + f_{i+2}). \tag{8.16}$$

The approximation (8.16) presents us with the well-known Simpson's rule of integration on the subinterval $[x_i, x_{i+2}]$.

Once the area a_i has been computed by the Simpson rule, the integral $I = \int_a^b f(x)dx$ is approximated as the sum of all integrals a_i,

$$\int_a^b f(x)dx \approx \frac{h}{3}\left[f_1 + 2\sum_{i=1}^{N/2-1} f_{2i+1} + 4\sum_{i=1}^{N/2} f_{2i} + f_{N+1} \right], \tag{8.17}$$

and we arrive at the composite Simpson's rule. It can immediately be seen from the formula (8.17) that the number N of grid sub-intervals must be even in order to apply the Simpson rule in the problem.

The above results can be further generalised as follows. Consider a polynomial $p_k(x)$ of degree k, where we require that $p_k(x_n) = f(x_n)$ for $n = i, i+1, \ldots, i+k$. In other words, we consider a polynomial passing through $k+1$ consecutive points where the function values are available. The area under the graph of the function $f(x)$ over a sub-interval $[x_i, x_{i+k}]$ is then approximated as

$$\int_{x_i}^{x_{i+k}} f(x)dx \approx a_i = \int_{x_i}^{x_{i+k}} p_k(x)dx,$$

and the resulting integral I is approximated by summation of all areas a_i.

Using local polynomials at each sub-interval $[x_i, x_{i+k}]$ with consecutive summation is known as the composite Newton-Cotes rules of numerical integration on regular grids (Davis and Rabinowitz 1975). The idea of interpolating the integrand function $f(x)$ by a polynomial $p_k(x)$ of degree k was pivotal in the development of the Newton-Cotes rules. The trapezoidal rule ($k = 1$) and the Simpson rule ($k = 2$) discussed above represent the first two rules in the Newton-Cotes family. They are, probably, the most well-known integration rules used in practical computations. The reason for their extensive use is twofold. Firstly, Newton-Cotes methods with $k > 2$ do not necessarily provide the most accurate estimate of the integral. For example, in the numerical integration problem considered in Davis and Rabinowitz (1975) increasing the polynomial degree k from $k = 2$ up to $k = 21$ resulted in a larger integration error in the latter case. Secondly, it is often difficult to apply a composite Newton-Cotes rule with $k > 2$ on a grid with an arbitrary number of grid sub-intervals, as the total number N of sub-intervals is required to be a multiple of k. That is why in many experimental applications the integral evaluation is restricted by the use of the composite trapezoidal rule (8.15) or the composite Simpson rule (8.17) and further in the text we consider the trapezoidal and Simpson rules only.

8.2.3 *Two-Dimensional Newton-Cotes Formulas*

Once the integration techniques have been understood in the one-dimensional $(1-d)$ case, they can be easily expanded to the two-dimensional $(2-d)$ case. Consider the unit square $D = [0,1] \times [0,1]$, where a regular grid is generated. Namely, let us consider a set of points $x_i, i = 1, \ldots, N+1$ on the interval $[0,1]$, where we require

that $x_1 = 0$, $x_{i+1} = x_i + h$, $i = 1, \ldots, N$, and the grid step size h is defined as $h = 1/N$. Similarly, we consider points y_j, $j = 1, \ldots, N+1$ on the interval $[0, 1]$ and generate a one-dimensional grid in the y-direction as $y_1 = 0, y_{j+1} = y_j + h, j = 1, \ldots, N$. The grid node position in the unit square is then given by (x_i, y_j) and we have a grid of square elements $c_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$.

A composite rule of integration in the $2-d$ case exploits the same idea as in the $1-d$ case. We have

$$I = \int_0^1 \int_0^1 f(x, y) dx dy = \sum_{i,j} I_{ij}, \tag{8.18}$$

where

$$I_{ij} = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} f(x, y) dx dy. \tag{8.19}$$

Hence, the integration problem is reduced to the integral evaluation in each subdomain c_{ij}. Integration on square elements can, in turn, be further reduced to consecutive application of the one-dimensional formulas. In other words, the integral (8.19) can be re-written as

$$I_{ij} = \int_{y_j}^{y_{j+1}} F(y) dy, \tag{8.20}$$

where

$$F(y) = \int_{x_i}^{x_{i+1}} f(x, y) dx.$$

We then employ $1-d$ Newton-Cotes formulas discussed in Sect. 8.2.2 in order to evaluate the function $F(y)$ in the square cell c_{ij}. Once the values of $F(y)$ have been computed, the same integration rule is applied to approximate the integral (8.20).

Different integration rules use different local approximation of the integrand $f(x, y)$ on a single grid cell c_{ij} (e.g., see Davis and Rabinowitz 1975). The simplest evaluation of the integral (8.19) can be done under the assumption that the function $f(x, y)$ is approximated by a constant on each grid cell. Such approximation results in the midpoint rule of integration:

$$I_{ij} \approx A_{ij} f(x_{i+1/2}, y_{j+1/2}), \tag{8.21}$$

where $A_{ij} = h^2$ and the node $(x_{i+1/2}, y_{j+1/2}) = (x_i + h/2, y_j + h/2)$ is the midpoint of the cell c_{ij}.

The trapezoidal rule of integration implies the approximation of $f(x, y)$ by a linear function on each sub-domain c_{ij}. Correspondingly, the integral I_{ij} is evaluated as

$$I_{ij} \approx \frac{h^2}{4}\left[f(x_i, y_j) + f(x_{i+1}, y_j) + f(x_i, y_{j+1}) + f(x_{i+1}, y_{j+1})\right]. \quad (8.22)$$

The Simpson rule of integration is a result of approximation of the integrand $f(x, y)$ by a quadratic polynomial in the square cell c_{ij}. The application of this rule in the cell c_{ij} requires that the data $f(x, y)$ are available at points (x_{i+q}, y_{j+r}), where $q = 0, 1, 2$ and $r = 0, 1, 2$. The function $f(x, y)$ is then integrated in the cell c_{ij} by the Simpson rule as

$$\begin{aligned} I_{ij} \approx \frac{h^2}{36}\big[&f(x_i, y_j) + f(x_i, y_{j+2}) + f(x_{i+2}, y_j) + f(x_{i+2}, y_{j+2}) \\ &+ 4\left(f(x_i, y_{j+1}) + f(x_{i+1}, y_j) + f(x_{i+2}, y_{j+1}) + f(x_{i+1}, y_{j+2})\right) \\ &+16 f(x_{i+1}, y_{j+1})\big]. \end{aligned} \quad (8.23)$$

Note that, like in the $1-d$ case, integration by the Simpson rule requires an even number N of grid sub-intervals in each direction x and y of a $2-d$ regular grid.

8.3 The Coarse Grid Problem

In this section we review a so called 'coarse grid' problem that was previously studied in detail in Petrovskaya and Petrovskii (2010), Petrovskaya et al. (2012), and Petrovskaya and Venturino (2011) because of its importance in ecological applications. For the sake of simplicity, the problem in this section is illustrated by $1-d$ examples, but the conclusions made in the $1-d$ case are also true for $2-d$ problems considered later in the text.

The coarse grid problem is closely related to the concept of the convergence rate discussed in Sect. 8.2. We know that, given the distance h between neighbouring points on a regular grid, the integration error (8.7) is controlled by the expression (8.9). However, we can only rely upon the error estimate (8.9) if the grid step size h is sufficiently small, i.e. if we deal with *fine* grids. Meanwhile if the distance h between grid nodes is not very small, it may happen that the error estimate (8.9) does not hold. In the latter case we cannot tell which integration method is better when we compare two integration methods based on their convergence rate (8.9). Correspondingly, a *coarse* grid is defined as a grid where one cannot apply the error estimate (8.9) to evaluate the integration error.

The above statement is illustrated in Fig. 8.3. We first consider the integrand

$$f(x) = \frac{e^x - 1}{e - 1}, \qquad x \in [0, 1], \quad (8.24)$$

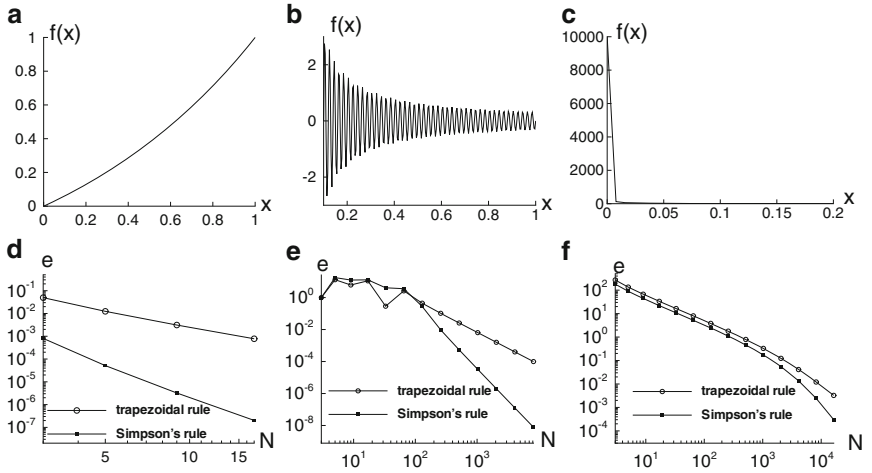

Fig. 8.3 The coarse grid problem: the comparison of the convergence rate for the trapezoidal rule (*solid line*, *open circle*) and the Simpson rule (*solid line*, *closed square*). (**a**) The integrand function (8.24) for $x \in [0, 1]$. (**d**) The convergence rate for the integrand (8.24) is as predicted by the error estimate (8.9). (**b**) The integrand function (8.25) over the interval $x \in [0.1, 1]$. (**e**) The convergence rate for the integrand (8.25). The error estimate (8.9) becomes true when the number N of grid sub-intervals is $N > N^* \approx 64$. (**c**) The integrand function (8.26). The function is shown at the sub-interval $[0, 0.2]$ for the sake of visualisation, while the integral is taken for $x \in [0, 1]$. (**f**) The convergence rate for the integrand (8.26). The error estimate (8.9) does not hold on coarse grids, unless at least one grid node is placed in the sub-region of the steep gradient

shown in Fig. 8.3a. The error graphs for the integrand (8.24) are shown in Fig. 8.3d, where the relative integration error (8.7) is computed for integration by the trapezoidal rule and the Simpson rule. It has been discussed in Sect. 8.2 that the convergence rate of the Simpson rule is much better than the convergence of the trapezoidal rule. Decreasing the grid step size from h_0 to $h_1 = h_0/2$ results in the error reduction $e(h_1) = (1/16)e(h_0)$ for the Simpson rule, while for the trapezoidal rule we have $e(h_1) = (1/4)e(h_0)$. It can be seen from Fig. 8.3d that in the case of the integrand function (8.24) this conclusion is true for any $N \geq 2$ considered in the problem.

Meanwhile, the above conclusion about the convergence rate does not hold for an arbitrary integrand function $f(x)$. Consider now a rapidly oscillating function

$$f(x) = \frac{\sin(100\pi x)}{\pi x}, \quad x \in [0.1, 1], \tag{8.25}$$

shown in Fig. 8.3b. The convergence of numerical integration methods, when the function (8.25) is integrated, is presented in Fig. 8.3e. The integration error on grids with $N \leq 64$ is very large for the both trapezoidal and Simpson's rule of integration. However, the most essential feature of the integration is that we cannot tell if the Simpson method is more accurate, unless we have the number of grid sub-intervals

$N > 64$, i.e. unless the grid step size becomes $h \approx 0.015$. The integration error of the Simpson rule remains approximately the same as the error of the trapezoidal rule on coarse grids with $N < N^* \approx 64$.

The coarse grid problem is further illustrated by consideration of the function

$$f(x) = \frac{1}{x + 0.0001}, \quad x \in [0, 1], \tag{8.26}$$

shown in Fig. 8.3c. It can be seen from the figure that the function (8.26) has a very narrow domain where the gradient is very steep. For the sake of illustration the function is shown on the sub-interval $x \in [0, 0.2]$, while the integration is carried out over the unit interval $x \in [0, 1]$. Our previous experience with the integrand (8.25) of Fig. 8.3b tells us that we can expect a big integration error when the number of grid nodes is not sufficient to resolve the domain of a steep gradient. This conclusion is confirmed by the convergence curve shown in Fig. 8.3f. The initial coarse grid with grid step size $h = 0.5$ cannot capture the sub-region of the steep gradient that has the width $w \sim 0.01$. Even when we make the grid step size smaller by halving each grid sub-interval, the whole sub-domain of the steep gradient remains 'invisible' to the integration method, as it is still located between two grid nodes where the function values are available. Hence both the trapezoidal and Simpson rules provide similar (and very inaccurate) results, unless we insert at least one grid node in the sub-region of the steep gradient. That happens when we have an unrealistically big number $N \approx 5{,}000$ of grid sub-intervals on a regular grid. Any grid with $N < N^*$ is a coarse grid where the error estimate (8.9) does not hold. Accordingly, any grid with $N > N^*$ is a fine grid where we can rely upon (8.9).

It was discussed in Petrovskaya and Petrovskii (2010) that the number N^* of grid subintervals when the grid becomes 'sufficiently refined', i.e.when we can rely upon the error estimate (8.9), can be evaluated from the knowledge of the shape of the integrand function. Let Δx be a characteristic width of a spatial heterogeneity described by a given integrand, e.g. the width of a single peak in (8.25). Then integration on a regular grid will give an inaccurate answer until at least one grid point falls into the heterogeneity region. We therefore have

$$N^* = s\frac{1}{\Delta x}, \tag{8.27}$$

where 1 in the numerator stands for the length of the domain of integration and $s \geq 1$ is a numerical coefficient depending on the type of the heterogeneity. If $f(x)$ is a monotone function on the interval $[x_1, x_2]$, then we consider the function values at two points, e.g., $x_1 + \delta$ and $x_2 - \delta$, where $0 < \delta < 0.5(x_1 + x_2)$, as the minimum 'amount of information' required to reconstruct $f(x)$ over $[x_1, x_2]$ as these data are sufficient for linear polynomial approximation of $f(x)$. Consequently, a sub-region of a steep gradient in (8.26) can be resolved by inserting into it just one grid point. Meanwhile, we need at least three grid points e.g., $x_1 + \delta$, x_2, and $x_3 - \delta$, to resolve a peak which spans the interval $[x_1, x_3]$, as that will result in linear approximation at each of subintervals $[x_1, x_2]$ and $[x_2, x_3]$ where $f(x)$ is a monotone function.

The most important conclusion that follows from the above consideration is that the grid coarseness should be evaluated in terms of the integration error rather than by the number of grid nodes. Hence the definition of a coarse grid depends strongly on the spatial pattern of the integrand function. It can immediately be seen from Fig. 8.3 that a grid considered as coarse for one integrand function can be a fine grid for another integrand. One way to improve very poor accuracy of integration on coarse grids would therefore be to use an irregular grid where most of the grid nodes would be concentrated in sub-regions that present difficulties in their numerical integration (i.e. peaks or sub-regions with a steep function gradient). Integration techniques on irregular grids are discussed in Sect. 8.6. However, in ecological applications it often is not possible to use irregular grids adapted to a spatial pattern of the density distribution because that pattern is usually not known a priori. On the other hand, coarse grids are widespread in ecological monitoring, as there are usually financial, ecological and other restrictions that do not allow for a big number of measurements and the data available in the problem are sparse. Thus the problem of accuracy control on coarse grids remains one of the most difficult problems in ecological monitoring and it is still far from being solved. We will discuss several particular examples of coarse grids in ecological applications further in the text.

8.4 Numerical Integration in Ecological Problems

In this section we consider the application of the methods reviewed in Sect. 8.2 to ecological monitoring and control. As we have already discussed in the introduction, one key problem of ecological monitoring is to obtain an accurate estimate $\tilde{I}(N)$ of the pest population size I in a given area under conditions when the population density is only known at $N + 1$ locations. It follows immediately from our study in Sect. 8.2 that the problem of evaluating the pest population size from discrete data can be considered as a problem of numerical integration. Indeed, installing traps in a domain where sampled data are collected and processing trap counts means that the discrete integrand function is defined at the nodes of a computational grid and methods (8.15) and (8.17) can be applied. However, several underlying assumptions should be made before we implement numerical integration rules in ecological problems.

8.4.1 Problem Statement and Underlying Assumptions

In our work we consider collecting information about a pest insect via trapping and we assume that a trapping procedure is done as described in the introduction. In this section we also assume that the traps are installed at the nodes of a regular grid, which is a common situation in ecological applications (Ferguson et al. 2000; Holland et al. 1999). Irregular grids will be discussed in Sect. 8.6.

As we already mentioned in the previous section, numerical integration techniques are essentially based on the underlying assumption that the integrand function is continuous. Meanwhile, if we consider an agricultural field, where pest insects are monitored, the distribution of the insects over the field is, of course, discrete. Hence, in order to apply numerical integration techniques in the problem we have to transform the discrete population distribution into a continuous function that we will refer to as "the population density". The population density can be obtained from the discrete distribution of the pest insects by allocating a certain area to each insect and assuming that only one insect can be found within that area at the fixed time t.

Another important underlying assumption is that the number of insects caught in each trap is an accurate representation of the absolute population density in its catchment area. The transformation techniques that allow one to link trap counts to the absolute density have been briefly discussed in the introduction. We also assume that the information about the population density at a given time and location can indeed be adequately obtained from trap counts, as depending on the biological and behavioural traits of the monitored species, the population density distribution can possibly change over the time of the traps' exposure. The spatial scale of variations in the population density distribution for walking insects usually sampled with pitfall traps is known to be 30–40 m (Holland et al. 1999). Meanwhile, typical dispersal distances for walking insects are estimated to be 1 m or less per day (Vinatier et al. 2010), which obviously corresponds to the spread area of the order of 1 m^2 per day. Hence the distance insects can move over 1 week (i.e. an average time of trap exposure) is $\sqrt{7} \approx 2.6$ m, which is about one order of magnitude less than the spatial scale of inherent variation. We conclude from the above that the spatial density distribution reconstructed from traps counts can approximately be relied upon as being static and methods of numerical integration can be applied.

Once the trap counts have been acquired, we can obtain the values of the population density at the nodes of a regular grid, i.e. at the trap locations, and we therefore can approximate the integral I, i.e. the total number of insects in the field, by a selected method of numerical integration. However, application of numerical integration in ecological monitoring and control is more difficult than a conventional integration problem. The standard numerical integration technique usually implies that a computational grid can be made sufficiently fine to provide the required accuracy. This requirement is not realistic in an ecological monitoring routine where the number of traps installed in a field cannot be made large. For example, a typical agricultural field in the United Kingdom has a characteristic size of the order of a few hundred meters. The number of traps installed over such a field very rarely exceed a few dozen (Blackshaw 1983; Ferguson et al. 2000; Holland et al. 1999). Moreover, we cannot increase the number of traps and repeat the trapping procedure if we are not happy with the accuracy of our original estimate, as a repeated trapping will inevitably be done under different conditions. Hence every time we compare the accuracy of several methods of numerical integration, we should keep in mind that in many ecological applications we may deal with numerical integration on coarse grids where the error estimate (8.9) cannot be applied. Further examples will be provided in the next sub-section.

8.4.2 Numerical Integration of Data Obtained from a Mathematical Model

Despite plenty of experimental data being available in the pest insect monitoring problem, we first apply our numerical integration techniques to the data obtained as a result of computer simulation. We use computer simulation for generating ecologically meaningful data because we want to subsequently increase the number of traps (i.e. the number of grid nodes) in order to investigate the integration error $e(N)$ for each integration method we employ in the problem. Thus we take our data from an ecologically sound mathematical model of population dynamics in order to be able to compute the function $e(N)$ for various N of our choice. Namely, we consider the spatially explicit predator-prey model with the Allee effect (Murray 1989; Turchin 2003). In dimensionless form the system is as follows:

$$\frac{\partial u(x,y,t)}{\partial t} = d\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + \beta u(u-b)(1-u) - \frac{uv}{1+\Lambda u}\,, \tag{8.28}$$

$$\frac{\partial v(x,y,t)}{\partial t} = d\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + \frac{uv}{1+\Lambda u} - mv\,, \tag{8.29}$$

where $x \in [0,1]$, $y \in [0,1]$, the functions $u(x,y,t)$ 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 (8.28–8.29) is solved numerically for a range of parameters, and the function $u(x,y,t)$ is then considered as the density of the pest insect in the problem. Solving the system (8.28–8.29) requires us to generate a regular spatial grid as described in Sect. 8.2. A discussion of the numerical solution along with the choice of the initial and boundary conditions has been provided in the paper Petrovskaya et al. (2012) where similar computer simulations have been done.

We begin our consideration from the simplest computational case where the pest insect population density is generated from the $1-d$ counterpart of the system (8.28–8.29). The parameters of the $1-d$ system of equations as well as the initial and boundary conditions required for its numerical solution are given in the paper Petrovskaya and Petrovskii (2010). The solution $u(x,t)$ of the $1-d$ system of equations is considered at fixed time $\hat{t}$ of our choice. We therefore have a $1-d$ spatial density distribution $u(x,\hat{t}) \equiv u(x)$ of the pest insect over the unit interval $x \in [0,1]$. This spatial density distribution is available to us at grid nodes x_i, $i = 1,2,\ldots,N+1$ only, but we can control the number $N+1$ of grid nodes in our computations by adding new nodes to a coarse grid or by removing them from a fine grid. Thus we first solve the $1-d$ system (8.28–8.29) on a grid with a very big number $N+1 = N_f + 1 = 2^{15}+1$ of nodes. Once the density distribution $u(x)$ has been obtained at fixed time $\hat{t}$ on an extremely fine grid of

$N_f + 1$ nodes, the integral $I = \int_0^1 u(x)dx$ is evaluated on that grid. The result $\tilde{I}(N_f)$ of that evaluation then is considered as the true integral I and is stored for further computations along with the spatial density $u(x)$ computed on a grid of $N_f + 1$ nodes. We then decrease the number of grid nodes and consider several approximations $\tilde{I}(N)$ for values $N \ll N_f$. Let us note that we do not re-compute the density function $u(x_i), i = 1, \ldots, N + 1$ every time that a new number N is chosen. The values of $u(x)$ are always taken from the 'exact solution' computed on a grid of $N_f + 1$ nodes at time $\hat{t}$, where we make a projection of the function $u(x_i), i = 1, \ldots, N_f + 1$ obtained on the fine grid to a coarse grid every time that we take a new, smaller number $N + 1$ of nodes. The details of this computational technique are provided in our previous work (Petrovskaya and Petrovskii 2010).

It is well-known (e.g., see Malchow et al. 2008; Petrovskii et al. 2003) that the properties of the spatial distribution $u(x)$ considered at a given time $\hat{t}$ are determined by the diffusion d. The density distribution can evolve into a monotone function if the diffusion d is of the order of 1 or larger. An example of a monotone density distribution is shown in Fig. 8.4a. For smaller values of $d \ll 1$ the initial conditions $u(x, 0),\ v(x, 0)$ evolve into an ensemble of irregular humps and hollows (see Fig. 8.4b), where the number of peaks gets bigger for smaller values of d. The density distributions from Fig. 8.4 present us with two somewhat extreme cases of ecologically meaningful integrand functions, while there can be one or two peaks in the domain for intermediate values of the diffusion coefficient d. Thus it is interesting to compare the accuracy of numerical integration for the two spatial patterns shown in the figure. Namely, we compare the results of the trapezoidal rule (8.15), the Simpson rule (8.17) and the results of the total population size evaluation by a statistical method.

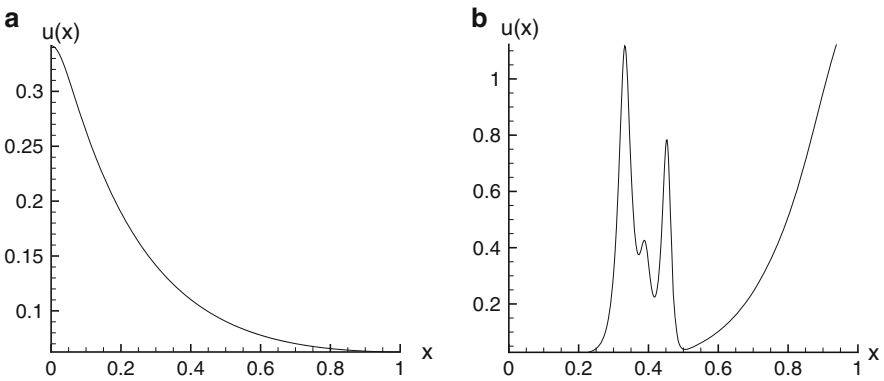

Fig. 8.4 Ecological test cases. (**a**) The spatial distribution of the pest population density $u(x)$ for the diffusivity $d = 10^{-4}$. Other parameters along with the initial and boundary conditions used to generate one-dimensional density distributions are discussed in Petrovskaya and Petrovskii (2010). (**b**) A 'multi-peak' density distribution obtained for the diffusion coefficient $d = 10^{-5}$

The statistical method commonly used in the evaluation of pest abundance is based on the computation of the sample mean pest population density (Davis 1994). The sample mean value $M(N)$ is given by a generic formula (8.1) where we have $K = N + 1$ in our problem. The expression (8.1) acts as an approximation to the true mean value. An approximation $\tilde{I}(N)$ to the actual pest population size I can then be found by multiplying the sample mean by the area of the field A, that is

$$I \approx \tilde{I}(N) = AM(N). \tag{8.30}$$

Consider the evaluation technique (8.30) in the $1-d$ case, so that the area A is given by the length $L = b - a$ of the interval $[a, b]$ where traps are installed. If the sampling positions x_i, $i = 1, \ldots, N + 1$, are equidistant, i.e. $x_{i+1} = x_i + h$ where $h > 0$ is constant, Eq. (8.30) can be written as

$$\tilde{I}(N) = \frac{L}{N+1} \sum_{i=1}^{N+1} u_i = \hat{h} \sum_{i=1}^{N+1} u_i = \sum_{i=1}^{N+1} \hat{h} u_i \approx \int_a^b u(x)dx, \tag{8.31}$$

where $u_i = u(x_i)$, $\hat{h} = L/(N + 1)$. It is readily seen that Eq. (8.31) coincides with the simplest method of numerical integration with weights $\omega_i = \hat{h}$. The convergence rate (8.9) of the integration rule (8.31) is $e = Ch$, where C is a constant (Davis and Rabinowitz 1975). Hence if the number of traps is big enough to resolve all features of the integrand function $u(x)$, the rule (8.31) should be inferior to more accurate integration methods such as the trapezoidal and Simpson rule.

The results of numerical integration of the density distributions shown in Fig. 8.4 are given in Table 8.1. It can be seen from the table that for the function $u(x)$ presented in Fig. 8.4a integration by the Simpson rule gives very accurate results even on a grid with a very small number of grid nodes. If we install three traps over the unit interval where the density measurements are made, evaluation of the total population size by the Simpson rule can be done with the error of 0.2 %, while the statistical rule provides us with an error over one hundred times bigger. Moreover, generally the error on each consecutive grid is smaller in comparison with the error on a previous grid.

Table 8.1 The relative integration error (8.7) for the $1-d$ density distributions of Fig. 8.4. The errors computed for the density distributions shown in Fig. 8.4a, b are marked with superscript (a) and (b), respectively. The first column gives the number $N + 1$ of the grid nodes. The error for each distribution (a) and (b) is computed by the statistical rule (8.31) (the column marked as e_{stat}), by the trapezoidal rule (8.15) (the column e_{TR}), and by the Simpson rule (8.17) (the column e_{SR})

$N+1$	$e_{stat}^{(a)}$	$e_{TR}^{(a)}$	$e_{SR}^{(a)}$	$e_{stat}^{(b)}$	$e_{TR}^{(b)}$	$e_{SR}^{(b)}$
3	0.283845	0.140895	0.002056	0.220479	0.056794	0.334066
5	0.132744	0.023493	0.015641	0.041373	0.245472	0.308365
9	0.064319	0.001141	0.006310	0.036879	0.138367	0.102665
17	0.033164	0.000373	0.000877	0.021570	0.025521	0.012094

Meanwhile, the situation is very different when we consider the density function shown in Fig. 8.4b. The application of the Simpson rule on grids with 3, 5 and 9 nodes does not have any advantage in comparison with the statistical rule. Clearly, in the case of the 'multi-peak' density distribution of Fig. 8.4b we have to deal with a 'coarse grid problem' where the density $u(x)$ is not well approximated on a grid with a small number of nodes. One important indicator of the coarse grid problem is that the error can oscillate between two grids, so that adding new nodes to the grid does not consistently make the error smaller until the integrand function is well resolved. An example is given by the error $e_{TR}^{(b)}$ of the trapezoidal rule on grids with $N = 3$ and $N = 5$ nodes.

It also is worth noting here that the statistical rule gives a more accurate answer on a grid of five nodes, while the error of the trapezoidal and the Simpson rule remains big on this grid. However, below we will see that even accurate results obtained on coarse grids are not reliable, as a slight change in the spatial pattern of the density function may result in a big jump in the integration error when the same numerical integration method is used in the problem.

8.4.3 Numerical Integration of $2-d$ Data

We now generate a $2-d$ density distribution $u(x, y)$ from numerical solution of the system of Eqs. (8.28–8.29). 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 (8.28–8.29) at any fixed time $\hat{t}$ provides us with the discrete density distribution $u_i \equiv u(x_i, y_j), i = 1, \ldots, N+1, j = 1, \ldots, N+1$, where grid nodes (x_i, y_j) are the points where traps are located. Similarly to the $1-d$ case we consider two density distributions whose spatial pattern is strongly different from each other. The density distribution shown in Fig. 8.5a presents a continuous front, while the density distribution of Fig. 8.5b is an example of a late stage of the patchy invasion (Petrovskii et al. 2005, 2002).

The computation carried out in the $1-d$ problem is repeated for the $2-d$ density distributions of Fig. 8.5. The results of numerical integration by the methods (8.30), (8.22) and (8.23) are shown in Table 8.2. It is readily seen from the table that the accuracy of integration depends again on the spatial pattern of the density function. Integration of the continuous front shown in Fig. 8.5a already gives a small integration error on grids with a small number of traps, the Simpson method being the most accurate method of integration. Let us recall that, in many ecological studies, a relative error of 100 % (i.e. $e(N) \sim 1$) is still regarded as acceptable, while the error $0.2 < e(N) < 0.5$ is considered as being good accuracy (Pascual and Kareiva 1996; Sherratt and Smith 2008). Hence numerical integration of the continuous front provides us with an accurate answer even on a grid with three grid nodes in each direction.

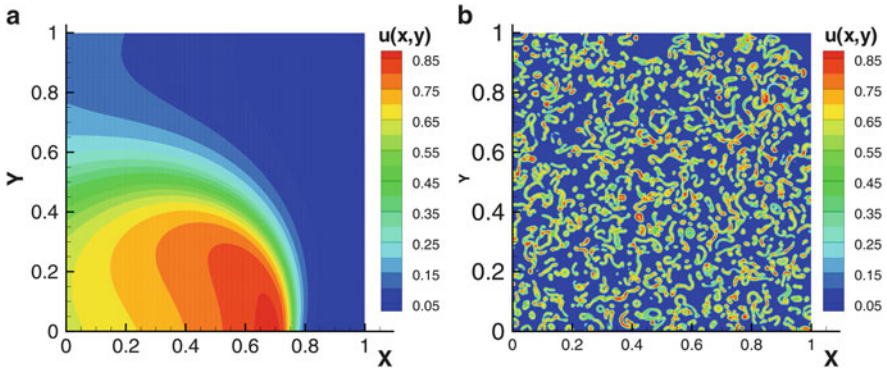

Fig. 8.5 Density function $u(x, y)$ as predicted by the population dynamics model (8.28–8.29) for different parameter values. (**a**) A snapshot of a continuous front. (**b**) A snapshot of the population density at a late stage of the patchy invasion

Table 8.2 The relative integration error (8.7) for the $2-d$ density distributions of Fig. 8.5. The errors computed for the density distribution shown in Fig. 8.5a, b are marked with superscript (a) and (b), respectively. The first column gives the number $N+1$ of the grid nodes in the direction x and y of a regular grid in the unit square. The error for each distribution (a) and (b) is computed by the statistical rule (8.31) (the column marked as e_{stat}), by the trapezoidal rule (8.22) (the column e_{TR}), and by the Simpson rule (8.23) (the column e_{SR})

$N+1$	$e_{stat}^{(a)}$	$e_{TR}^{(a)}$	$e_{SR}^{(a)}$	$e_{stat}^{(b)}$	$e_{TR}^{(b)}$	$e_{SR}^{(b)}$
3	0.1383	0.0506536	0.0255829	0.421591	0.496434	0.492878
5	0.064104	0.0142134	0.0221032	0.179808	0.263172	0.179825
9	0.032304	6.51693e-004	0.00389531	0.112412	0.111526	0.067423
17	0.017627	2.86861e-004	9.55669e-005	0.086713	0.064729	0.053797

Meanwhile, the more complex spatial structure of the density distribution of Fig. 8.5b requires a bigger number of grid nodes to provide the same level of accuracy. Moreover, on analysing the performance of the Simpson rule (8.23) on grids with $N+1<9$, we see that it is not more accurate than the other methods on coarse grids. On grids where the spatial pattern of the density function $u(x, y)$ is not well resolved it is hard to say which method is more accurate. This conclusion is further confirmed by numerous computations of approximate integrals made for various spatial distributions in the paper Petrovskaya et al. (2012).

8.4.4 Examples of Numerical Integration of Field Data

In this section we apply numerical integration techniques to field data of ecological monitoring. The aim of this study is to check what can be the smallest number of grid nodes (i.e. the number of traps in the agricultural field) used for accurate

evaluation of the pest insect population size, given the spatial distribution of the pest insect density. Obviously, density measurements made under real-life conditions cannot provide us with the data on a very fine grid. However, some experimental data contain information sufficient to extract a sequence of grids with a smaller number of nodes from the original grid and to compare the results of numerical integration on grids with various numbers of nodes.

We first illustrate our approach by considering data that have already been used in our earlier paper (Petrovskaya et al. 2012) where numerical integration techniques have been applied to experimental data collected for a New Zealand flatworm population (*Arthurdendyus triangulatus*) by Murchie and Harrison (2004). The data on flatworm abundance at different locations were collected by means of trapping where the traps were positioned at the nodes of a 12×12 regular grid. Spacing between two traps was 2 m in each direction. The traps were examined every week and the numbers of flatworms caught were counted. The various 12×12 grid trap systems caught 465–748 flatworms per sampling period (Murchie and Harrison 2004). Other details of the trapping procedure can be found in Murchie and Harrison (2004) and Petrovskaya et al. (2012).

Two examples of the density distributions obtained from trap counts are shown in Fig. 8.6. The trap counts have been linked to the local population density $u(x, y)$ by dividing the trap counts at each location by 4 m^2 (i.e. by the area of the grid cell) (Byers et al. 1989; Raworth and Choi 2001). For the sake of numerical integration, we have then extracted a sub-grid with $N + 1 = 11$ traps in each direction from the trap data originally collected in the field. This has been done because integration by the Simpson rule requires an odd number $N + 1$ of grid nodes. Having integrated the

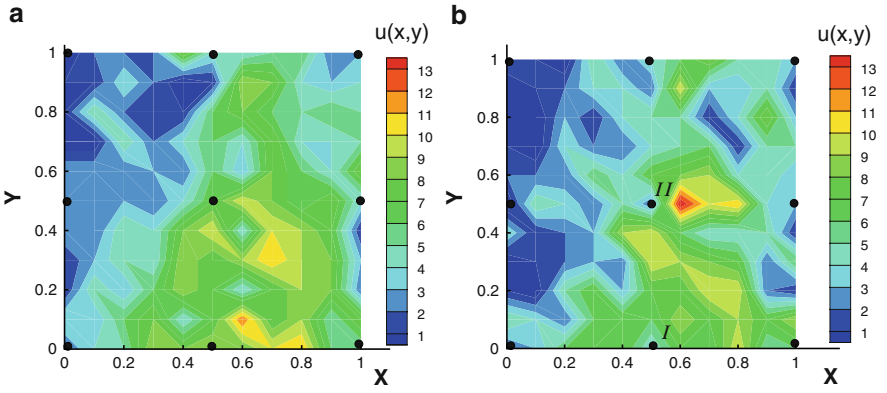

Fig. 8.6 Numerical integration of field data on a coarse grid of nine nodes. The nodes of a regular grid are shown as closed circles in the figure. The field data present flatworm spatial distributions over the study area (see Murchie and Harrison 2004; Petrovskaya et al. 2012 for more details).(**a**) Numerical integration of the density distribution gives good accuracy even on a grid with a very small number of nodes. (**b**) Two grid nodes (node *I* and node *II* in the figure) fall into small patches of different density on a coarse grid. Since the density values at those locations are not representative, numerical integration on a coarse grid results in a big integration error

Table 8.3 The approximation of the total population size and the integration error on a regular grid of 3×3 nodes for the field data taken from the paper Petrovskaya et al. (2012). The approximate integral is computed by the statistical rule (8.30) (the column I_{stat}), by the trapezoidal rule (8.22) (the column I_{TR}) and by the Simpson rule (8.23) (the column I_{SR}). The rows marked (a) and (b) in the table correspond to the density distributions shown in Fig. 8.6a, b, respectively

case	I	I_{stat}	e_{stat}	I_{TR}	e_{TR}	I_{SR}	e_{SR}
(a)	611	411	0.327	488	0.202	561	0.082
(b)	544	289	0.469	269	0.506	247	0.545

population density over the fine grid of $11 \times 11 = 121$ nodes, we have reproduced the total number I of collected insects. This number is further considered as the exact value of the population size.

Let us now compute the population size on a regular grid of $3 \times 3 = 9$ nodes and compare the population size obtained by numerical integration over this grid with the value I obtained for the density distributions shown in Fig. 8.6a and Fig. 8.6b on the original grid of 11×11 nodes. The 9 traps on a coarse grid are stationed as shown in Fig. 8.6 and we take the density values at those locations from the original grid.

The results of numerical integration on a grid of nine nodes are presented in Table 8.3. We compute the integral by the statistical rule (8.30) (the column I_{stat}), by the trapezoidal rule (8.22) (the column I_{TR}) and by the Simpson rule (8.23) (the column I_{SR}). We also compute the relative integration error (8.7) for each of the rules above (the columns marked as e_{stat}, e_{TR}, e_{SR}, respectively). The exact population size (i.e. the integral computed on a grid of 121 nodes) is $I^{(a)} = 611$ for the density distribution shown in Fig. 8.6a and $I^{(b)} = 544$ for the density distribution shown in Fig. 8.6b.

The results presented in the table confirm our previous conclusion that the accuracy of evaluation depends heavily on the spatial pattern of the density function $u(x, y)$. It is seen from the table that the integration of the density distribution shown in Fig. 8.6a gives good accuracy even on a grid with the number of nodes as small as nine nodes. This result lead us to the conclusion that robust information about the population size of pest insect population can be obtained using far fewer traps per unit area, provided that the spatial density pattern is not very patchy.

In the case of the density distribution in Fig. 8.6b some information about the density function $u(x, y)$ has been lost, as two grid nodes have fallen into small sub-regions (patches) where the density is strongly different from the density in the surrounding domain; see nodes *I* and *II* in the figure. The density values at those nodes made a misleading contribution to the sum (8.4) and that resulted in a big integration error. Meanwhile it is worth noting that even in the case (b) the relative error of the population size estimate still remains smaller than 55 % and such accuracy can still be considered as acceptable for large scale monitoring programmes (Northing 2009).

Let us also note that we had to transform the grid of 12×12 traps to conduct our computational study, as application of the Simpson method was not possible on a grid with an even number of nodes in each direction. Meanwhile, we would like to

Table 8.4 An example of trap count data for *Pterostichus melanarius* obtained by trapping with pitfall traps (The data are taken from the paper Alexander et al. (2005))

5	0	1	2	4	1	1	38	5	4	3	3	1	13	5	6
7	13	1	0	1	0	0	12	2	0	1	8	12	10	1	0
6	3	0	0	4	2	1	1	2	3	5	11	12	11	3	0
2	5	1	7	8	6	15	0	3	1	0	6	2	8	1	0
7	5	1	2	0	2	0	0	4	3	3	0	9	7	4	1
3	7	6	0	0	1	6	0	5	2	0	2	16	13	6	2
4	6	3	0	5	8	1	4	3	6	2	26	11	1	5	2
2	2	2	7	9	5	13	5	3	14	26	42	9	15	1	4
1	0	3	2	11	0	3	7	8	11	14	22	24	5	5	0
6	1	4	16	15	11	0	11	12	13	16	20	12	7	5	4
1	0	4	1	11	2	11	7	6	6	0	3	4	6	0	0
9	6	3	2	7	7	6	8	11	25	18	9	2	1	2	1
3	2	6	15	5	18	24	4	8	16	6	11	6	1	0	0
3	7	3	22	27	34	0	41	21	37	16	10	3	7	2	3
12	12	30	25	23	15	19	12	6	9	9	4	10	6	3	6
11	7	11	26	38	19	16	19	11	13	13	0	5	4	2	10

emphasise that application of so called higher order numerical integration methods can be made on grids with an arbitrary number of grid nodes in each direction. A numerical integration method that would have the same convergence rate as the Simpson method could be designed for the original grid of 12×12 nodes. The application of such a method, however, would be a much more difficult technical task and it is beyond the scope of our paper. Thus we only provide a brief discussion of more general methods of numerical integration in Sect. 8.6.

Our conclusion about the accuracy of numerical integration is further illustrated by another set of field data taken from the paper Alexander et al. (2005). The trap counts for beetles *Pterostichus melanarius* obtained by trapping with pitfall traps are presented in Table 8.4. Field sampling for data in the table was performed on a 16×16 regular grid of traps installed in a conventionally managed 4 ha winter wheat field in Devon, UK; see Alexander et al. (2005) for more details. The density distribution obtained from Table 8.4 is shown in Fig. 8.7a. Again, for the purpose of our study we have to transform the original grid of 16×16 nodes into a grid where the Simpson method of integration can be applied. In the case of Table 8.4 we found it more convenient to augment the table rather than extract a grid with a smaller number of nodes. Generating a 17×17 computational grid from the original data should allow us to compute the integration error on a sequence of regular grids of 3×3, 5×5 and 9×9 nodes and to compare the accuracy of integration on those grids. Thus we added another row and column to the original grid. Hypothetical data for that addition were generated under the requirement to keep the same structure of the spatial pattern as in the original density distribution (see Fig. 8.7b). The value of the integral (i.e. using the total number of trap counts) on the new grid of 17×17 nodes is $I = 1{,}980$.

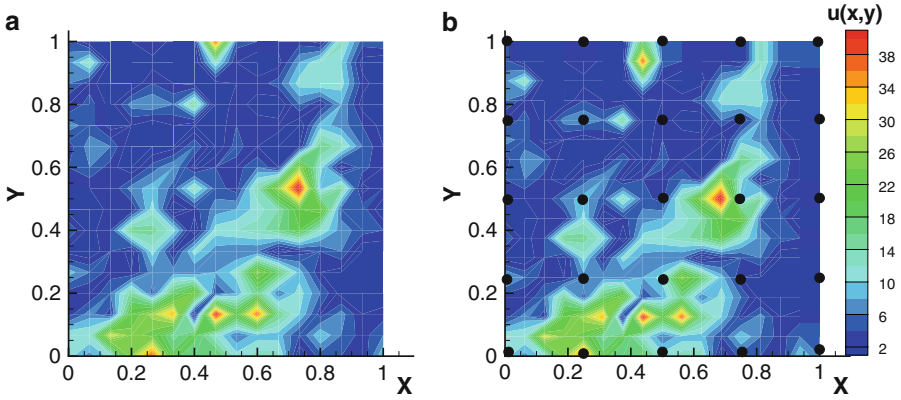

Fig. 8.7 The spatial density distribution obtained from the trap counts in Table 8.4. (**a**) The density function $u(x, y)$ based on the original data in Alexander et al. (2005). (**b**) Hypothetical data have been added to the original table in order to generate a 17×17 regular grid. The data have been generated to preserve the spatial structure of the original density distribution. An example of a regular grid (5×5 nodes) on which integral is computed is shown as a set of closed circles in the figure

It can be seen from Fig. 8.7b that the spatial pattern of the density distribution is similar to the spatial pattern of the $1 - d$ function (8.26) studied in Sect. 8.3. The density distribution is mostly homogeneous (cf. the function (8.26) on the interval $x \in [0.01, 1]$) with several small patches where the density is very high (cf. the function (8.26) for $x \in [0, 0.01]$). From the study of the convergence graph for the function (8.26) we predict that grids with 3×3, 5×5 and 9×9 nodes should be considered as coarse grids for the density distribution $u(x, y)$ of Fig. 8.7b, as small patches of the high density are not resolved on those grids. Hence the Simpson method will not have a visible advantage over the other integration methods employed in the problem. On the other hand, a big sub-domain where the density is almost homogeneous will be already well-resolved on coarse grids and integration over that sub-domain should give us an accurate contribution to the integral over the whole domain. Meanwhile, the number of patches with high density is not big and the density localised there is only approximately 30 times bigger than the density in the homogeneous sub-domain, while this ratio is approximately 5,000 for the function (8.26). Thus we expect a reasonably small integration error on coarse grids.

An example of a regular computational grid (a grid of 5×5 nodes) used in our computation is shown in Fig. 8.7b. The location of the nodes on a this grid confirms our analysis in the previous paragraph. Namely, all small patches of high density fall in between the grid nodes, but the density values on grid nodes are already representative enough to give accurate information about the density function in sub-domains where the density is an almost homogeneous function (see also Table 8.4). The results of numerical integration are shown in Table 8.5. The relative error is within 35 % even on a grid with three traps in each direction. However, as

Table 8.5 The approximate integral and the relative integration error (8.7) for the density distribution of Fig. 8.7b on a sequence of regular grids. The first column gives the number $N+1$ of the grid nodes in the direction x and y of a regular grid. The approximate integral and the error is computed on each grid by the statistical rule (8.30) (the columns marked as I_{stat} and e_{stat} in the table), by the trapezoidal rule (8.22) (the columns I_{TR} and e_{TR}), and by the Simpson rule (8.23) (the columns I_{SR} and e_{SR})

$N+1$	I_{stat}	e_{stat}	I_{TR}	e_{TR}	I_{SR}	e_{SR}
3	1,507	0.239	1,344	0.321	2,332	0.178
5	1,679	0.152	1,632	0.175	1,561	0.212
9	1,659	0.162	1,692	0.145	1,748	0.117

predicted, increasing the number of traps from three to nine in each direction does not significantly increase the accuracy of integration methods because small patches are still not resolved. Also, the Simpson method is not definitively superior to the statistical method and the trapezoidal rule, as the convergence rate (8.9) does not hold on coarse grids.

One important conclusion drawn from our consideration of $1-d$ and $2-d$ ecological distributions is that the accuracy of an estimation depends strongly on how the pest insects are dispersed across the agricultural field. The question of accuracy has been the focus of ecological research for a long time (Dent 2000; Vlug and Paul 1986; Ward et al. 1985). Reliable recommendations have been provided on the minimum number of traps required for obtaining an accurate estimate of a particular pest insect species based on the assumption that the pest insect density distribution is close to homogeneous (Binns et al. 2000; Karandinos 1976; Southwood and Henderson 2000). This assumption is true for many species, but as we could see in this section there also exist many ecologically important cases where the pest density is heterogeneous and can be aggregated into several patches (see also Barclay 1992; Ferguson et al. 2003). In the latter case we can anticipate an inaccurate estimate of the total pest population size, as a relatively small number of traps normally used in the trapping procedure may not be sufficient to resolve highly localised sub-domains of non-zero density. In the next section we discuss an extreme case of a single density patch in order to demonstrate that a conceptually different approach should be applied to evaluate the total population size for such spatial patterns.

8.5 Highly Aggregated Density Distributions

In this section we consider highly aggregated density distributions that we also refer to as peak functions. Namely, we discuss spatial patterns where 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. Such distributions have ecological significance as they present at an early stage of the biological invasion

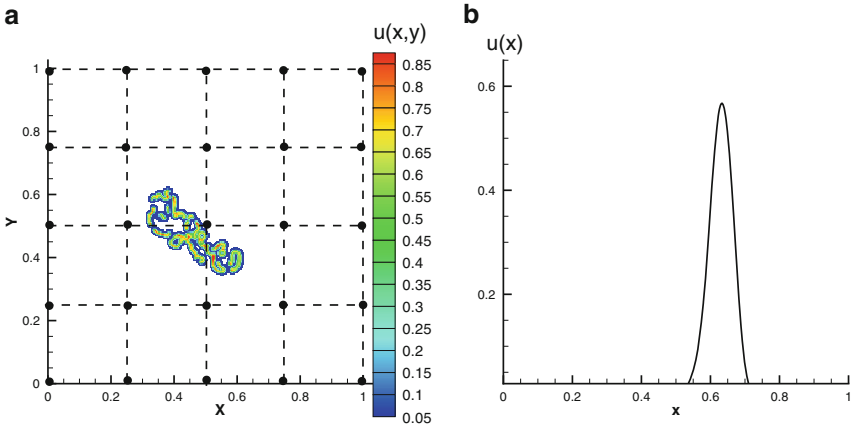

Fig. 8.8 (**a**) The pest population density distribution $u(x, y)$ at an early stage of patchy invasion. The highly aggregated density function $u(x, y)$ has been obtained from numerical solution of Eqs. (8.28–8.29). The traps used to measure the density $u(x, y)$ are installed at the nodes of a regular coarse grid as shown in the figure. (**b**) A one-dimensional counterpart of the density distribution of Fig. 8.9a

(Shigesada and Kawasaki 1997). It is clear that timely and accurate evaluation of the total number of pest insects at the beginning of biological invasion is beneficial for the cultivation of the agricultural product. At the same time the application of numerical integration methods to highly aggregated density distributions is a very difficult task, as the exact location of the high density sub-domain is normally not known in the problem. Thus, instead of installing the traps locally in the patch of the non-zero density in order to increase the accuracy of integration, traps have to be stationed at the nodes of a regular grid over the entire domain where monitoring is made. That, in turn, may result in the most unfavourable situation when the entire patch of non-zero density falls in between grid nodes.

Examples of highly aggregated density distributions are depicted in Fig. 8.8 where the density function was modelled by solving Eqs. (8.28–8.29) in the $2-d$ case (see Fig. 8.8a) and in the $1-d$ case (see Fig. 8.8b). It can be seen from Fig. 8.8a that the sub-region of non-zero density is entirely missed on a coarse grid of 5×5 nodes and we should significantly increase the number of nodes in order to resolve that sub-region. Given natural limitations on the number of traps that present in ecological applications, two basic questions arise. The first question is: What is the minimum number $N_t + 1$ of traps required to achieve desirable accuracy if a highly aggregated density distribution is numerically integrated? Also, we have to answer the related question: What can be an alternative measure of accuracy on a regular grid of traps where $N < N_t$?

The answer to the questions above were offered in the paper Petrovskaya and Embleton (2013). It has been shown there that a standard methodology does not work when the density of a highly aggregated pest population is measured by a trapping procedure with a small number $N + 1$ of traps installed. The uncertainty

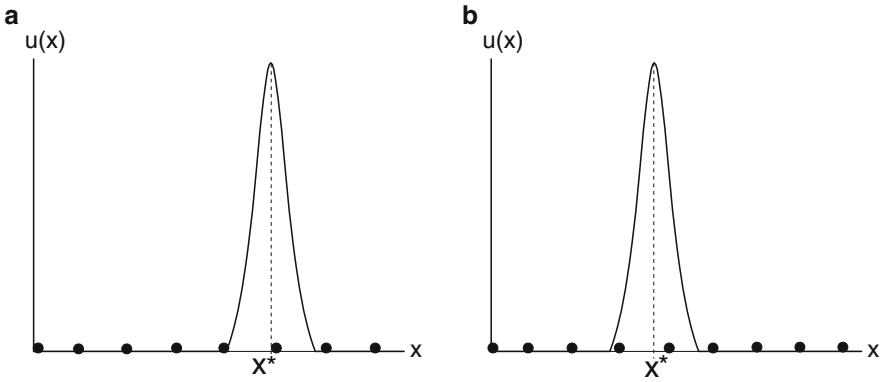

Fig. 8.9 Numerical integration of a highly aggregated density distribution. The accuracy of integration depends on the peak location x^* with respect to the nodes of a regular grid. (**a**) The peak sub-domain contains only one grid node. (**b**) The same peak is now located in a different region, so that two grid nodes lie within the peak sub-domain when the same regular grid is generated

in measurements made on coarse grids is so strong that an estimate $\tilde{I}(N)$ of the integral I becomes a random variable. As a result, the integration error also becomes a random variable with a high magnitude and we cannot control the accuracy of evaluation. In other words, we cannot consider the condition (8.11) for peak functions on grids with small N, as, depending on the peak location, we sometimes will obtain a very accurate answer and sometimes our answer on the same regular grid will be well beyond the accuracy range. An example illustrating this statement is shown in Fig. 8.9, where we have one grid node within the peak sub-domain in Fig. 8.9a. If we move the peak on the same grid, so that the location of the maximum x^* becomes different, two grid nodes will fall into the peak region (see Fig. 8.9b). As a result, the peak function will be better resolved and we will have a more accurate estimate of the integral.

Since the integration error is considered as a random variable on coarse grids where a location of the density patch is not known to us, it was therefore suggested in Petrovskaya and Embleton (2013) that we have to compute the probability of achieving an integration error within a certain accuracy range instead of computing the error itself. Namely, we compute the probability $p(h)$ (or $p(N)$ in some cases) that the condition (8.11) holds. The probability $p(h)$ is then considered an alternative measure of accuracy when we integrate a high aggregation density distribution on a regular grid with a small number of nodes.

Grids, where the integration error becomes a random variable because of the insufficient information about the integrand function and where we have to compute the probability of an accurate evaluation of the integral are referred to as *ultra-coarse* grids (Embleton and Petrovskaya 2013; Petrovskaya and Embleton 2013; Petrovskaya et al. 2012). It is clear that if we keep increasing the number of nodes on a regular grid, then sooner or later we are able to integrate the peak function with very good accuracy. We therefore have the threshold number N_t of

grid sub-intervals, where the desirable accuracy of pest population size evaluation cannot be guaranteed for any $N < N_t$. An immediate consequence of this result is that an estimate of the pest population size per se becomes unreliable if the number N of traps in the field is $N < N_t$.

The above results are illustrated by a simple example of a $1-d$ peak function. Consider the following density distribution (the Lorentzian) on the unit interval $x \in [0, 1]$,

$$u(x) = \begin{cases} \dfrac{\delta^2}{4}\dfrac{1}{4(x-x^*)^2+\delta^2/4} - \dfrac{1}{5}, & x \in [x^*-\delta/2, x^*+\delta/2], \\ 0, & otherwise, \end{cases} \tag{8.32}$$

where δ is the peak width and x^* is the location of the maximum point. Let us emphasise again that the location x^* is not known to us, so the peak can be located at any point of the sub-interval $[\delta/2, 1-\delta/2]$. The peak function (8.32) is shown in Fig. 8.10a for the peak width $\delta = 0.06$.

Consider now numerical integration of the function (8.32) by the trapezoidal rule (8.15). Let a regular grid of $N+1$ nodes be generated in the domain $[0, 1]$ as $x_1 = 0, x_{i+1} = x_i + h, i = 1, \ldots, N$, where the grid step size is $h = 1/N$. For the purpose of our study we require that the grid step size is $h > \delta/2$. We start from $h = 0.25$ (i.e., five equidistant grid nodes over the unit interval) and decrease h by adding new nodes to the grid until the grid step size is so small that the condition $h > \delta/2$ is broken. For each grid step size h we compute the probability $p(h)$ of getting an accurate estimate of the integral, provided that the peak is arbitrarily located in the domain. The accuracy we impose in the problem is $e(N) \le \tau_0 = 0.25$, where $e(N)$ is the relative integration error (8.7). The details of the computation of $p(h)$ can be found in Petrovskaya and Embleton (2013).

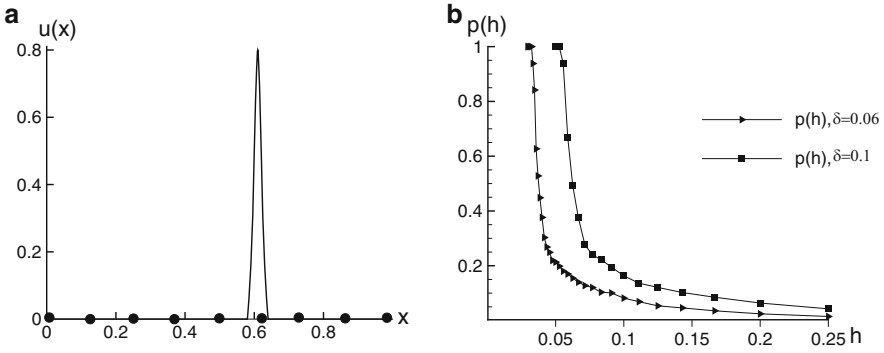

Fig. 8.10 (**a**) The peak function (8.32) with the peak width $\delta = 0.06$ on a regular grid of 9 nodes. (**b**) The probability curves $p(h)$ obtained for the function (8.32) with the peak width $\delta = 0.06$ (*solid line*, *closed right triangle*) and $\delta = 0.1$ (*solid line*, *closed square*)

Two probability curves $p(h)$ are shown in Fig. 8.10b for the function (8.32) with the peak width $\delta = 0.06$ and $\delta = 0.1$. It can be seen from the figure that for each probability curve there exists the threshold value h_t such that $p(h) = 1$ for any $h < h_t$. It was shown in Petrovskaya and Embleton (2013) that the grid step size h_t for which the error (8.7) becomes deterministic, that is $p(h_t) = 1$ and $p(h) < 1$ for any $h > h_t$, can be evaluated as

$$h_t = \alpha_t \delta, \tag{8.33}$$

where δ is the peak width and α_t is constant for any given tolerance τ in the accuracy condition (8.11). The value of α_t was computed in Petrovskaya and Embleton (2013) as $\alpha_t \approx 0.81$ for $\tau = 0.25$.

Let us, for example, integrate the function (8.32) with the peak width $\delta = 0.1$ on a regular grid with the grid step size $h = 0.1$ (i.e., a grid of 11 nodes). We have $h > h_t \approx 0.08$ and it follows from Fig. 8.10b that our chance $p(h)$ to evaluate the integral within the accuracy range $e(N) < 0.25$ is $p(h) \approx 0.2 = 20\,\%$. In other words, there is an 80 % chance that the error of our evaluation will be bigger than $\tau = 0.25$ when we evaluate the pest abundance for the peak function (8.32) on a regular grid of 11 nodes. Consider now a grid with $h = 0.07$ (15 grid nodes). Since the distance between nodes is now $h < h_t$, we will always have the error $e(N)$ of integral evaluation smaller than 0.25, no matter where the peak is located. The probability $p(h)$ of getting the error within the accuracy range $e(N) < 0.25$ is $p(h) = 1$.

In the $1 - d$ case the grid step size h is given by $h = 1/N$ and we can therefore evaluate the minimum number $N_t = 1/h_t$, such that the desirable accuracy of integration is guaranteed on a grid of $N_t + 1$ nodes. Furthermore, it has been discussed in Petrovskaya et al. (2013) that in ecological problems the width δ of the highly aggregated density distribution can be written as

$$\delta = \omega\sqrt{d}, \tag{8.34}$$

where d is the diffusion coefficient. Another coefficient ω in the expression (8.34) depends on the system's parameters. It was shown in Petrovskii and Malchow (2001) and Petrovskii et al. (2003) that the value ω is relatively robust to changes in the parameter values and can typically be considered as $\omega \approx 25$. Hence the threshold number N_t can be evaluated as

$$N_t = \frac{1}{\alpha_t \delta} \approx \frac{1}{\alpha_t \omega \sqrt{d}}. \tag{8.35}$$

For example, the ecologically meaningful density distribution of Fig. 8.8b was generated for the diffusion $d = 10^{-4}$. The estimate (8.35) gives us the grid step size as $h_t \approx 0.2$ and the corresponding number of grid nodes is $N_t + 1 \approx 6$.

Understanding accuracy requirements for highly aggregated density distributions is important when a sampling plan is designed for pest insect monitoring and control. As we already mentioned in the introduction, a standard procedure of the

risk evaluation in pest management is to compare an estimate of the total number of pest insects with a certain critical number and to make a decision based on that comparison. We discussed in the previous sections that the error in the estimation of pest abundance becomes worse as the number of samples decreases (see also Binns et al. 2000). However, consideration of the extreme case of a random error brings into the problem another risk factor related to the uncertainty in integral evaluation when the number $N + 1$ of traps is small. Taking this risk factor into account may constitute an important task in the whole process of designing an appropriate methodology for decision making in pest insect management.

8.6 Evaluating Pest Abundance on Irregular Grids

So far we have considered using methods of numerical integration to evaluate pest population abundance when the sampling plan is a regular grid, i.e. the samples are taken at regular spatial intervals. However, it may be that an irregular grid is prescribed in a pest monitoring programme. Furthermore, even if a regular grid has been selected as the intended sampling plan, taking samples at precisely regular intervals may not be possible in practise. The landscape of an agricultural field may have natural obstacles (e.g. a bush or a tree) that will make trap installation on the nodes of a regular grid impossible. One or many of the samples may then have to be taken at a location shifted from that which was intended due to an obstruction of some kind, hence the resulting grid of samples is irregular. We thus now investigate the accuracy of numerical integration methods formulated on an irregular grid. Our analysis is focused on a $1-d$ problem for the sake of simplicity, but, as in previous sections, our results can be readily extended to a $2-d$ problem.

8.6.1 Generation of Irregular Grids

We consider several types of grids with varying degrees of irregularity: a slightly irregular grid, a quasi-random grid, and a random grid. We use the term 'slightly irregular' to refer to a simple example of an irregular grid, whereby a single sampling location is shifted from the position prescribed by a regular sampling plan. We generate such a grid by first constructing a regular grid as was explained in Sect. 8.2. A single interior node x_i, for some $i = 2, \ldots, N$ is then perturbed according to the following transformation:

$$x_i^{irreg} = x_i + h\left(r - \frac{1}{2}\right), \tag{8.36}$$

where x_i is a node location on a regular grid, and $r \in (0, 1)$ is a uniformly distributed random variable. The transformation (8.36) is further illustrated in Fig. 8.11b.

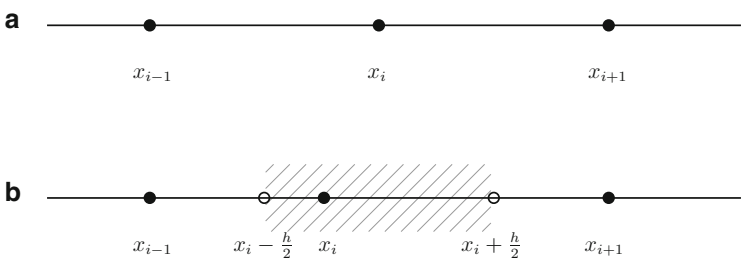

Fig. 8.11 (**a**) An interior grid node x_i for some $i = 2, \ldots, N$ is a fixed distance h from its neighbouring grid nodes in accordance with a regular sampling plan. (**b**) An interior grid node x_i^{irreg} which has been perturbed according to the transformation (8.36) (the superscript is omitted in the figure to make it consistent with Fig. 8.11a). The shaded region shows the possible locations for x_i^{irreg}, where this node is no longer an equal distance from its neighbouring nodes

A quasi-random grid has an increased level of irregularity whilst preserving some structure. Such grids are generated in a similar way to the method discussed above for the slightly irregular grids. The difference is that instead of a single interior node being perturbed, *all* interior nodes are perturbed. That is, the transformation (8.36) is applied to all interior nodes $x_i, i = 2, \ldots N$ of the regular grid. This form of grid is closely related to the so called 'centric systematic' sampling plan (e.g. see Milne 1959) whereby the field is divided into sections and a sample is taken from a random location within each section. Our version differs only in that we have fixed the boundary points so as to preserve the interval of integration as $[a, b]$.

A random sampling plan is often viewed favourably from a theoretical viewpoint as it is considered to avoid introducing bias into the estimate (Bliss 1941; Legg and Moon 1994; Reisen and Lothrop 1999; Silver 2008), the concern being that a systematic distribution of samples will somehow coincide with the distribution of the pests. We therefore take into consideration such a distribution of samples in our investigation and generate the points $x_i, i = 1, \ldots, N + 1$ as follows:

$$x_i = a + r(b - a), \qquad i = 1, \ldots, N + 1, \tag{8.37}$$

where $r \in (0, 1)$ is a uniformly distributed random variable. The points $x_i, i = 1 \ldots N + 1$ are then sorted into ascending order and the endpoints on a random grid are then replaced as

$$x_1 = a, \qquad x_{N+1} = b. \tag{8.38}$$

An example of a random grid is shown in Table 8.6. The grid of nine nodes presented in the table was generated over the interval $[0, \pi]$ using a standard function *rand*() in Visual C++.

Table 8.6 An example of random grid over the interval $[0, \pi]$

i	1	2	3	4	5	6	7	8	9
x_i	0.0	0.816881	1.05838	1.43716	1.489	1.58434	1.70697	1.74214	3.14159

8.6.2 Numerical Integration on Irregular Grids

We now look at the accuracy of pest abundance estimates obtained by methods of numerical integration on the grids outlined above. We will be using the statistical rule, the trapezoidal rule and Simpson's rule to evaluate the pest abundance. Since the statistical rule (8.31) has no spatial dependence it can be applied to regular and irregular grids alike.

Meanwhile, we must use different forms of the trapezoidal and Simpson's rules to those which have been mentioned above in order to be able to apply them to irregular grids. The Newton-Cotes formulas can, of course, be applied in the case that the integrand function $f(x)$ is defined on the nodes of an irregular grid. The idea remains the same: replace the integrand by a polynomial function and integrate the polynomial instead. However, we cannot use formulas (8.15) and (8.17) designed for regular grids and we have to take into account a grid's irregularity when the weight coefficients ω_i are computed.

The trapezoidal rule on irregular grids is given by

$$I \approx \tilde{I} = \sum_{i=1}^{N} h_i \frac{(f_i + f_{i+1})}{2}, \tag{8.39}$$

where N is the number of grid sub-intervals, and the grid step size $h_i = x_{i+1} - x_i$ is variable rather than fixed as in the formula for regular grids. We use the following adapted version of Simpson's rule to handle irregular grids

$$I \approx \tilde{I} = \sum_{i=1}^{\frac{N}{2}} \frac{h_{2i-1} + h_{2i}}{6} \left(f_{2i-1} + 4f_{2i} + f_{2i+1}\right), \tag{8.40}$$

which also relies on the variable grid step size $h_i = x_{i+1} - x_i$. As with the conventional Simpson's rule (8.17), the number of grid nodes $N + 1$ is required to be odd.

We illustrate the convergence on irregular grids by considering a sequence of grids, where each grid is generated according to the relevant procedure outlined above. The number of grid subintervals is set on the first grid in the sequence as $N = N_0$, an estimate is obtained by means of a chosen numerical integration method and the relative error (8.7) is calculated. The number of grid sub-intervals is then increased to $N_1 = 2N_0$, a new grid is generated, and the estimate and subsequent relative error is recalculated. This process is repeated until the number of grid sub-intervals reaches some chosen value $N = N_{final}$. In the case of the

slightly irregular grids, we want to determine how perturbing a single node affects the convergence rate of a method of numerical integration, rather than how the position of the grid node which is perturbed affects the accuracy. As such, in each generation of the slightly irregular grids, the same interior grid node is perturbed. We will begin all of our calculations on a grid of three grid nodes which has only one interior node. The unperturbed position of this node lies at $x = (a + b)/2$, therefore, it will always be this central node which is perturbed in the generation of each slightly irregular grid. For grids with a more significant level of irregularity i.e. the quasi-regular and random grids, each grid generation is repeated a total of n_r times thus providing n_r values of the error for any given grid of $N + 1$ nodes. The mean error on a grid of $N + 1$ nodes is then calculated as

$$\mu(e) = \frac{1}{n_r} \sum_{i=1}^{n_r} e_i . \tag{8.41}$$

We first consider a standard mathematical test case where the integral of the function (8.10) is evaluated over a sequence of increasingly refined irregular grids according to the procedure outlined above. For the slightly irregular grids, the corresponding relative errors are shown in Table 8.7. It can be seen from the table that very little difference is made to the accuracy by perturbing a single node as the results for the regular and slightly regular grids are close to each other.

For the random grids, the mean of $n_r = 10^4$ evaluations of the error have been plotted in Fig. 8.12. The convergence rate of errors calculated over increasingly refined regular grids has also been plotted in each graph for comparison purposes (see dashed line in the figure). Random perturbation of the interior nodes affects the convergence rate with varying degrees of prominence depending on the method of numerical integration employed as can be seen in the figure. The behaviour of the convergence curve for the statistical rule shown in Fig. 8.12a is different from the convergence for the trapezoidal rule (Fig. 8.12b) and the Simpson rule (Fig. 8.12c), as the convergence rate of the method (8.31) on irregular grids is slower in comparison with the convergence on regular grids. Meanwhile the randomness introduced to the computational grid causes the convergence curves

Table 8.7 The relative integration error (8.7) for the function (8.10) on slightly irregular grids where the central node is shifted from its position on the original regular grid. The first column gives the number $N + 1$ of grid nodes. The error (8.7) is computed on an irregular grid (marked as the superscript "irreg" in the table) and compared with the corresponding error on a regular grid (the superscript "reg"). The error is computed for the statistical rule (8.31) (the columns marked as e_{stat} in the table), the trapezoidal rule (8.39) (the columns e_{TR}), and the Simpson rule (8.40) (the columns e_{SR})

$N+1$	e_{stat}^{reg}	e_{stat}^{irreg}	e_{TR}^{reg}	e_{TR}^{irreg}	e_{SR}^{reg}	e_{SR}^{irreg}
3	4.764e−001	4.829e−001	2.146e−001	2.163e−001	4.720e−002	4.681e−002
5	2.416e−001	2.423e−001	5.194e−002	5.409e−002	2.280e−003	1.786e−003
9	1.226e−001	1.227e−001	1.288e−002	1.288e−002	1.346e−004	1.342e−004
17	6.185e−002	6.185e−002	3.215e−003	3.222e−003	8.296e−006	8.295e−006

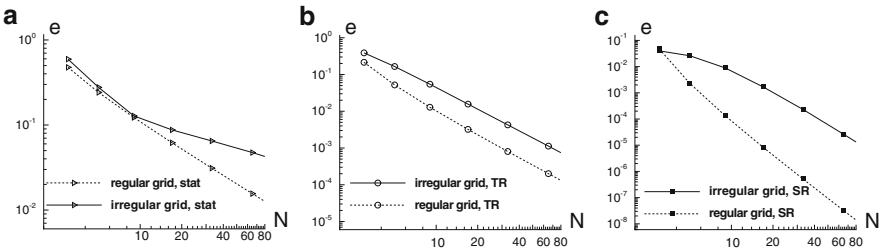

Fig. 8.12 Numerical integration of the function (8.10) on random grids. (**a**) The relative integration error (8.7) for the statistical rule (8.31) (*solid line*, *right open triangle*). The convergence curve is compared with the convergence on regular grids (*dashed line*, *right open triangle*). (**b**) Convergence curves for the trapezoidal rule (8.39) on random grids (*solid line*, *open circle*). The convergence curve for the method (8.15) on regular grids is shown as a *dashed line* in the figure. (**c**) Convergence curves for the Simpson rule (8.40) on random grids (*solid line*, *closed square*) and for the method (8.17) on regular grids (*dashed line*, *closed square*)

of the trapezoidal and Simpson's rules to be shifted upwards, that is, the resulting estimates are less accurate although they begin to converge at a similar rate to those formulated on regular grids as N increases. The higher the degree of the method applied, the more prominent the effect seems to be, although it should be noted that on average the accuracy still improves when a higher degree method is used.

8.6.3 Integration of Ecological Data on Irregular Grids

Let us now consider the accuracy of the numerical integration of ecologically significant data. Since we are required to perform repeated calculations over increasingly refined grids, we use simulated data as suitable field data is difficult to obtain. As earlier explained the simulated ecological population density functions were obtained through numerical solution of the $1-d$ system (8.28–8.29) on an extremely fine, regular grid of $N_f+1 = 2^{15}+1$ nodes on the interval $[a,b] = [0,1]$. Since the density functions are thus discrete rather than continuous, the method for generating the slightly irregular computational grid is now different to that outlined above although the fundamental ideas are the same.

We have available a fine grid of points $x_i^f, i = 1,\ldots,N_f+1$ where

$$x_1^f = a = 0, \qquad x_i^f = x_{i-1} + \frac{b-a}{N_f}, \quad i = 2,\ldots,N, \qquad x_{N_f+1}^f = b = 1.$$

To generate a slightly irregular grid of $N+1$ nodes, a regular grid is first obtained by extracting the required $N+1$ nodes from the available fine grid as

$$x_i = x_j^f, \qquad j = 1+(i-1)\left(\frac{N_f}{N}\right), \quad i = 1,\ldots,N+1. \tag{8.42}$$

A single interior node must then be perturbed, however, it must be perturbed to a value for which the population density is available. This is achieved by replacing an interior grid node as

$$x_i = x^f_{j+r}, \qquad r \in \left[-\frac{N_f}{2N}, \frac{N_f}{2N}\right] \tag{8.43}$$

for some $i = 2, \ldots, N$, where j is as given in (8.42) and r is a uniformly distributed random integer.

The generation of quasi-random grids for use with simulated ecological data is as follows. The endpoints are fixed as

$$x_1 = x^f_1, \qquad x_{N+1} = x^f_{N_f+1}, \tag{8.44}$$

and the interior points are defined as

$$x_i = x^f_{j+r}, \qquad r \in \left[-\frac{N_f}{2N}, \frac{N_f}{2N} - 1\right], \qquad i = 2, \ldots, N. \tag{8.45}$$

Note that here the upper limit of the interval to which r belongs is one less than that in (8.43) so as to avoid any nodes coinciding.

To extract a random grid from the available data, the grid nodes of the fine grid $x^f_i, i = 1, \ldots, N_f + 1$ are first permuted randomly. We shall denote the resulting points as $\tilde{x}^f_i, i = 1, \ldots, N_f + 1$. We begin to form a random grid of $N + 1$ nodes by selecting the first $N + 1$ nodes from the permuted fine grid so we have

$$x_i = \tilde{x}^f_i, \qquad i = 1, \ldots, N + 1. \tag{8.46}$$

The nodes $x_i, i = 1, \ldots, N + 1$ are then sorted into ascending order and the endpoints are replaced as

$$x_1 = a = 0, \qquad x_{N+1} = b = 1. \tag{8.47}$$

Let us now consider the three-peak simulated ecological test case as shown in Fig. 8.4b. As above, we generate a sequence of increasingly refined grids and the relative errors are calculated according to (8.7). It should be noted that since the exact value of the integral is not available to us for such discrete data, we have taken the approximation obtained by applying the trapezoidal rule to the extremely fine, regular grid of $N_f + 1$ nodes to be the 'exact' value of the pest abundance I. For the quasi-random and random grids, $n_r = 10^4$ of each grid are generated and the mean of the errors is calculated.

Convergence curves for the slightly irregular grids, where one node is randomly shifted from its original location on a regular grid, are shown in Fig. 8.13. The integration error (8.7) computed for the statistical rule (8.31) is presented in

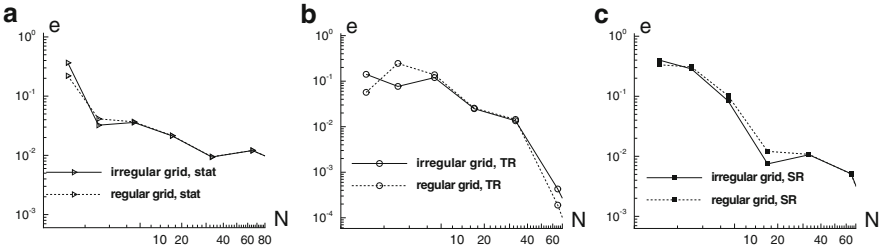

Fig. 8.13 Convergence curves on slightly irregular grids for the ecologically meaningful density distribution of Fig. 8.4b. Convergence on a sequence of grids where a central grid node is randomly shifted is compared to the convergence on regular grids. The figure legend is the same as in Fig. 8.12. (**a**) The statistical rule (8.31), (**b**) the trapezoidal rule (8.39), and (**c**) the Simpson rule (8.40) is implemented

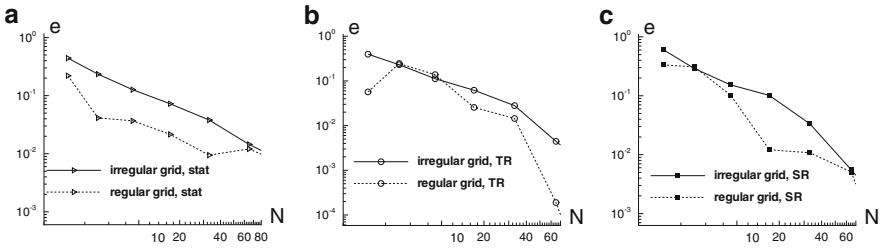

Fig. 8.14 Convergence curves on quasi-random grids for the ecologically meaningful density distribution of Fig. 8.4b. Convergence on a sequence of grids where each interior grid node is randomly shifted around its position on a regular grid is compared to the convergence on regular grids. The figure legend is the same as in Fig. 8.12. (**a**) The statistical rule (8.31), (**b**) the trapezoidal rule (8.39), and (**c**) the Simpson rule (8.40) is implemented

Fig. 8.13a, while the error for the trapezoidal rule (8.39) and the Simpson rule (8.40) is shown in Fig. 8.13b, c, respectively. The convergence results in the figure confirm our previous conclusion made for the function (8.10). A slight perturbation of grid regularity results in a slight perturbation in the integration error.

Let us now make a stronger perturbation of a regular grid and consider numerical integration on a sequence of quasi-random grids where each interior grid node is randomly shifted around its position on a regular grid. The corresponding convergence curves are shown in Fig. 8.14, where the figure legend is the same as in Fig. 8.12. It can be seen from the figure that increasing the degree of grid randomness in the problem results in a bigger integration error, no matter what integration method is used. This conclusion is further illustrated by consideration of the integration error on truly random grids; see Fig. 8.15. Again, the convergence curves shown in Fig. 8.15 for integration on regular grids always lie below convergence curves obtained for random grids for any integration rule employed in the problem.

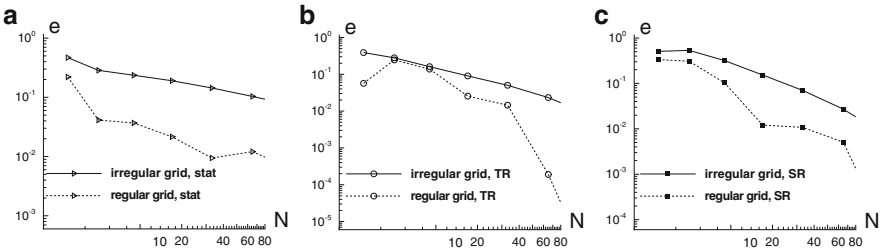

Fig. 8.15 Convergence curves on regular and random grids for the ecologically meaningful density distribution of Fig. 8.4b. The figure legend is the same as in Fig. 8.12. (**a**) The statistical rule (8.31), (**b**) the trapezoidal rule (8.39), and (**c**) the Simpson rule (8.40) is implemented

The results of our study demonstrate that grid randomisation leads to a bigger integration error on coarse and fine grids alike. Surprisingly, this conclusion is true even for the statistical method which has no spatial dependence. While further careful study of this issue is required, our first experience with the problem of numerical integration on random grids demonstrates that an equidistant distribution of traps is better than a random distribution.

8.7 Concluding Remarks

We considered the application of methods of numerical integration to the problem of evaluating pest insect abundance. Methods of numerical integration are well known and documented in the literature, but, to our best knowledge, they have never been applied in ecological problems. Meanwhile, employing advanced numerical integration techniques can be beneficial in the evaluation of total pest population size, as those techniques can help to improve the accuracy of evaluation. In our paper we studied a trapping procedure in an agricultural field and discussed how information about the pest population density at trap locations can be transformed into a numerical integration problem. However, our conclusions about the applicability of methods of numerical integration in ecological problems are general enough and therefore remain valid when the information about the local species density is obtained by another sampling technique.

The key idea behind numerical integration methods considered in the paper is to locally replace the existing density distribution by an approximated density distribution described by a polynomial function. The most straightforward way to apply numerical integration is to install traps at the nodes of a regular grid, but similar techniques can be designed for a random distribution of traps over an agricultural field. From a numerical integration viewpoint the method (8.31) widely used in ecological applications can be loosely interpreted as local approximation of the density function by a constant. While such approximation provides in some

cases rather poor accuracy, approximation by higher order polynomials (e.g. by a quadratic function) should result, according to the theory of numerical integration, in more accurate evaluation of pest abundance. It has been shown in the paper that advanced numerical integration techniques (e.g., the Simpson rule on regular grids) often provide a significantly more accurate estimate of the population size from trap data than the standard statistical approach (8.31). In many cases methods remain effective even when the distribution exhibit a complex spatial structure.

At the same time, it was discussed in the paper that the application of numerical integration methods in ecological problems may be restricted by the poor resolution of the density distribution on coarse grids. Our study demonstrated that numerical integration methods may become unreliable when pest abundance is evaluated from a heterogeneous density pattern on a coarse grid. For example, the accuracy of the Simpson method (8.17) is superior to the statistical rule (8.31) and the trapezoidal rule (8.15), but the Simpson method has no visible advantage over less accurate methods (8.31) and (8.15) when a strongly heterogeneous density distribution is considered on a coarse grid of traps. In the extreme case when the total population is localised in a small sub-domain, an estimate of the total population size becomes a random variable, and we cannot even tell whether or not the estimate is within a given accuracy range.

The coarse grid problem remains, in our opinion, the main obstacle to the implementation of numerical integration methods in IPM programmes. It was shown in the paper that grid coarseness is not defined by the number of traps available in the problem. For any fixed number of traps, that number can be considered as a grid with good resolution for one density pattern, while the same grid of traps can appear as a coarse grid, where the accuracy of evaluation is poor, for another density distribution. Our study confirmed that grid coarseness is directly related to the degree of heterogeneity, highly aggregated density distributions being the most difficult case for numerical integration. Meanwhile, 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). Thus an important conclusion that stems from our results is that any information about the spatial pattern of the pest insect density distribution must be used to its fullest extent (cf. Perry 1996; Perry and Hewitt 1991) in order to decide whether or not we can expect to obtain an accurate estimate of pest abundance. This conclusion is true for any numerical integration technique including the method (8.31), as examples studied in the paper reveal that an estimate of the mean density on coarse grids can be very far away from its true value. Let us also note that the unreliability of results on coarse grids should, in our opinion, be taken into account as another risk factor when a sampling plan is designed, and results of large-scale ecological monitoring should be interpreted accordingly.

Summarising the above, the main recommendation from our study is to implement methods of numerical integration that are based on approximation of the density distribution by higher order polynomials (e.g. the Simpson method). If heterogeneity in a spatial pattern is not well resolved, no method of numerical integration has an advantage over the other methods, as all of them will give equally

unreliable results. However, as soon as the heterogeneity is resolved, approximation of the density distribution by higher order polynomials will provide a more accurate estimate of the total pest population size.

Our study leaves a number of open questions. The most difficult and crucial issue is, of course, the question of how to get information about a spatial pattern of the density distribution in order to be able to predict the accuracy of integration. Another important issue related to the question above is the optimisation of trap locations. Grid adaptation to the spatial pattern can be made if we have the information about patches of high density. Numerical integration on an adapted grid should result in an improvement in accuracy, but its application requires further discussion of the technical details. Also, numerical integration techniques can be extended to domains of arbitrary shape, but resolution of a curvilinear boundary remains a topic for future research.

Finally, we would like to emphasise that numerical integration techniques still have to be validated for a broad variety of ecological test cases before they can be routinely used in ecological monitoring and control. However, identification and clear understanding of all theoretical aspects of numerical integration techniques can accelerate and simplify further incorporation of those techniques into IPM programmes and the issues that have been in the focus of this paper are important milestones along the way.

References

Alavanja MCR, Ross MK, Bonner MR (2013) Increased cancer burden among pesticide applicators and others due to pesticide exposure. CA Cancer J Clin 63:120–142

Alexander CJ, Holland JM, Winde L et al (2005) Performance of sampling strategies in the presence of known spatial patterns. Ann Appl Biol 146:361–370

Alyokhin A, Baker M, Mota-Sanchez D et al (2008) Colorado potato beetle resistance to insecticides. Am J Potato Res 85:395–413

Apostol TM (1974) Mathematical analysis. Addison-Wesley, Reading

Ausden M (1996) Invertebrates. In: Sutherland WJ (ed) Ecological census techniques: a handbook. Cambridge University Press, Cambridge

Barclay HJ (1992) Modelling the effects of population aggregation on the efficiency of insect pest control. Res Popul Ecol 34:131–141

Bates SL, Zhao JZ, Roush RT, Shelton AM (2005) Insect resistance management in GM crops: past, present and future. Nat Biotechnol 23:57–62

Binns MR, Nyrop JP, Van Der Werf W.(2000) Sampling and monitoring in crop protection: the theoretical basis for designing practical decision guides. CABI Publishing, Wallingford

Birmingham AL, Kovacs E, Lafontaine JP et al (2011) A new trap and lure for Drosophila melanogaster (Diptera: Drosophilidae). J Econ Entomol 104:1018–1023

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

Bliss CI (1941) Statistical problems in estimating populations of Japanese beetle larvae. J Econ Entomol 34:221–232

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

Burn AJ, Coaker TH, Jepson PC (1987) Integrated pest management. Academic, New York

Byers JA, 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
Christou P, Capell T, Kohli A et al (2006) Recent developments and future prospects in insect pest control in transgenic crops. Trends Plant Sci 11:302–308
Comins HN, Hassell MP, May RM (1992) The spatial dynamics of host-parasitoid systems. J Anim Ecol 61:735–748
Davis PM (1994) Statistics for describing populations. In: Pedigo LP, Buntin GD (eds) Handbook of sampling methods for arthropods in agriculture. CRC Press, Boca Raton, pp 33–54
Davis PJ, Rabinowitz P (1975) Methods of numerical integration. Academic, New York
Dent D (2000) Insect pest management. CABI Publishing, Wallingford
Embleton NL, Petrovskaya NB (2013) On numerical uncertainty in evaluation of pest population size. Ecol Complex 14:117–131
Embleton NL, Petrovskaya NB (2014) A novel approach to evaluation of pest insect abundance in the presence of noise. Bull Math Biol 76:718–743. doi:10.1007/s11538-014-9940-z
Ester A, van Rozen K (2005) Monitoring and control of Agriotes lineatus and A. obscurus in arable crops in the Netherlands. IOBC Bull Insect Pathog Insect Parasit Nematodes Melolontha 28:81–86
Evans EW, Rogers RA, Opfermann DJ (1983) Sampling grasshoppers (Orthoptera: Acrididae) on burned and unburned tallgrass prairie: night trapping vs. sweeping. Environ Entomol 12: 1449–1454
Ferguson AW, Klukowski Z, Walczak B et al (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. Ent Exp Appl 95:161–171
Ferguson AW, Klukowski Z, Walczak B et al (2003) Spatial distribution of pest insects in oilseed rape: implications for integrated pest management. Agric Ecosyst Environ 95:509–521
Gatehouse AMR, Ferry N, Edwards MG, Bell HA (2011) Insect-resistant biotech crops and their impacts on beneficial arthropods. Philos Trans R Soc B Biol Sci 366:1438–1452
Gwinner J, Harnisch R, Mück O (1996) Manual of the prevention of post-harvest grain losses. GTZ, Eschborn
Hagler JR, Jackson CG (2001) Methods for marking insects: current techniques and future prospects. Ann Rev Entomol 46:511–543
Higley LG, Pedigo LP (1996) Economic thresholds for integrated pest management. University of Nebraska Press, Lincoln
Hokkanen HMT (1991) Trap cropping in pest management. Ann Rev Entomol 36:119–138
Holland JM, Perry JN, Winder L (1999) The within-field spatial and temporal distribution of arthropods in winter wheat. Bull Entomol Res 89:499–513
Hutchins SH (1994) Techniques for sampling arthropods in integrated pest management. In: Pedigo LP, Buntin GD (eds) Handbook of sampling methods for arthropods in agriculture. CRC, Boca Raton, pp 73–97
Jepson PC, Thacker JRM (1990) Analysis of the spatial component of pesticide side-effects on non-target invertebrate populations and its relevance to hazard analysis. Funct Ecol 4:349–355
Karandinos MG (1976) Optimum sample size and comments on some published formulae. Bull Entomol Soc Am 22:417–421
Kogan M (1998) Integrated pest management: historical perspectives and contemporary developments. Ann Rev Entomol 43:243–270
Legg DE, Moon RD (1994) Bias and variability in statistical estimates. In: Pedigo LP, Buntin GD (eds) Handbook of sampling methods for arthropods in agriculture. CRC, Boca Raton, pp 55–69
Liebman M, Dyck E (1993) Crop rotation and intercropping strategies for weed management. Ecol Appl 3:92–122
Louws FJ, Rivard CL, Kubota C (2010) Grafting fruiting vegetables to manage soilborne pathogens, foliar pathogens, arthropods and weeds. Scientia Horticulturae 127:127–146
Malchow H, Petrovskii SV, Venturino E (2008) Spatiotemporal patterns in ecology and epidemiology: theory, models, and simulations. Chapman & Hall/CRC, London

Mayor JG, Davies MH (1976) A survey of leatherjacket populations in south-west England, 1963–1974. Plant Pathol 25:121–128
Metcalf RL, Luckmann WH (eds) (1982) Introduction to insect pest management. Wiley, London
Milne A (1959) The centric systematic area-sample treated as a random sample. Biometrics 15:270–297
Murchie AK, Harrison AJ (2004) Mark-recapture of 'New Zealand flatworms' in grassland in Northern Ireland. In: Proceedings crop protection in Nothern Britain: 2004, Association for Crop Protection in Northern Britain, Dundee, Scotland
Murray JD (1989) Mathematical biology. Springer, Berlin
Northing P (2009) Extensive field based aphid monitoring as an information tool for the UK seed potato industry. Aspects Appl Biol 94:31–34
Oerke EC (2006) Crop losses to pests. J Agric Sci 144:31–43
Okubo A (1986) Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. Adv Biophys 22:1–94
Pascual MA, Kareiva P (1996) Predicting the outcome of competition using experimental data: Maximum likelihood and Bayesian approaches. Ecology 77:337–349
Pedigo LP, Rice ME (2009) Entomology and pest management. Pearson Prentice Hall, Upper Saddle River
Perry JN (1996) Simulating spatial patterns of counts in agriculture and ecology. Comput Electron Agric 15:93–109
Perry JN, Hewitt M (1991) A new index of aggregation for animal counts. Biometrics 47: 1505–1518
Petrovskaya NB, Embleton NL (2013) Evaluation of peak functions on ultra-coarse grids. Proc R Soc A 469:20120665. http://dx.doi.org/10.1098/rspa.2012.0665
Petrovskaya NB, Embleton NL, Petrovskii SV (2013) Numerical study of pest population size at various diffusion rates. In: Lewis MA et al (eds) Dispersal, individual movement and spatial ecology. Lecture notes in mathematics. Springer, Berlin/Heidelberg 2071:355–386
Petrovskaya NB, Petrovskii SV (2010) The coarse-grid problem in ecological monitoring. Proc R Soc A 466:2933–2953
Petrovskaya NB, Petrovskii SV, Murchie AK (2012) Challenges of ecological monitoring: estimating population abundance from sparse trap counts. J R Soc Interface 9:420–435
Petrovskaya NB, Venturino E (2011) Numerical integration of sparsely sampled data. Simul Model Pract Theory 19:1860–1872
Petrovskii SV, Bearup D, Ahmed DA, Blackshaw RP (2012) Estimating insect population density from trap counts. Ecol Complex 10:69–82
Petrovskii SV, Li B-L, Malchow H (2003) Quantification of the spatial aspect of chaotic dynamics in biological and chemical systems. Bull Math Biol 65:425–446
Petrovskii SV, Malchow H (2001) Spatio-temporal chaos in an ecological community as a response to unfavorable environmental changes. Adv Complex Syst 4:227–250
Petrovskii SV, Malchow H, Hilker FM, Venturino E (2005) Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control. Biol Invasions 7:771–793
Petrovskii SV, Morozov AY, Venturino E (2002) Allee effect makes possible patchy invasion in a predator-prey system. Ecol Lett 5:345–352
Pimentel D (1995) Amounts of pesticides reaching target pests: environmental impacts and ethics. J Agric Environ Ethics 8:17–29
Pimentel D (ed) (1997)Techniques for reducing pesticide use: economic and environmental benefits. Wiley, New York
Pimentel D (2009) Pesticides and pest control. In: Peshin R, Dhawan AK (eds) Integrated pest management: innovation-development process, vol 1. Springer, Berlin, pp 83–87
Pimentel D, Greiner A (1997) Environmental and socio-economic costs of pesticide use. In: Pimentel D (ed) Techniques for reducing pesticide use: economic and environmental benefits. Wiley, New York, pp 51–78
Pimentel D, Pimentel M (2008) Food, energy and society. CRC, Boca Raton

Raworth DA, Choi MJ (2001) Determining numbers of active carabid beetles per unit area from pitfall-trap data. Ent Exp Appl 98:95–108
Reisen WK, Lothrop HD (1999) Effects of sampling design on the estimation of adult mosquito abundance. J Am Mosq Control Assoc 15:105–114
Ruberson JR (ed) (1999) Handbook of pest management. Marcel Dekker, New York
Sherratt JA, Smith M (2008) Periodic travelling waves in cyclic populations: field studies and reaction diffusion models. J R Soc Interface 5:483–505
Shoffner AV, Tooker JF (2013) The potential of genotypically diverse cultivar mixtures to moderate aphid populations in wheat (Triticum aestivum L.) Arthropod-Plant Interact 7:33–43
Shigesada N, Kawasaki K (1997) Biological invasions: theory and practice. Oxford University Press, Oxford
Silver JB (2008) Mosquito ecology: field sampling methods. Springer, New York
Smigocki AC, Ivic-Haymes S, Li H, Savić J (2013) Pest protection conferred by a beta vulgaris serine proteinase inhibitor gene. PLoS ONE 8:e57303. doi:10.1371/journal.pone.0057303
Snedecor GW, Cochran WG (1980) Statistical methods. The Iowa State Iniversity Press, Ames
Sohrabi F, Shishehbor P, Saber M, Mosaddegh MS (2013) Lethal and sublethal effects of imidacloprid and buprofezin on the sweetpotato whitefly parasitoid Eretmocerus mundus (Hymenoptera: Aphelinidae). Crop Protection 45:98–103
Southwood TRE, Henderson PA (2000) Ecological methods. Blackwell Science, Oxford
Stern VM (1973) Economic thresholds. Ann Rev Entomol 18:259–280
Stern VM, Smith RF, van den Bosch R, Hagen KS (1959) The integration of chemical and biological control of the spotted alfalfa aphid. Part I. The integrated control concept. Hilgardia 29:81–101
Taboada A, Pérez-Aguirre C, Assmann T (2012) A new method for collecting agile tiger beetles by live pitfall trapping. Ent Exp Appl 145:82–87
Turchin P (2003) Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press, Princeton
Vinatier F, Chailleux A, Duyck P-F et al (2010) Radiotelemetry unravels movements of a walking insect species in heterogeneous environments. Anim Behav 80:221–229
Vlug HJ, Paul H (1986) Sampling leatherjackets. Med Fac Landbouww Rijksuniv Gent 51:939–942
Ward SA, Rabbinge R, Mantel WP (1985) The use of incidence counts for estimation of aphid populations. 1. Minimum sample size for required accuracy. Neth J Plant Pathol 91:93–99