

3.2 Dispersal and dispersion in space

S.A. Ward, P.S. Wagenmakers and R. Rabbinge

3.2.1 Introduction

The models and techniques presented in Section 3.1 can be used to simulate the development of pest populations or disease epidemics during the course of a year, but they include no spatial component. It was assumed implicitly that the pests or diseases are distributed evenly among the sampling units (e.g., wheat tillers, potato leaves, etc.). This assumption, however, is rarely or never true: most pests or diseases are found in clusters or patches.

The effects of such aggregation can be extremely important, especially when the population's growth rate is influenced by local density-dependent factors such as predation or intra-specific competition for food or sites (Section 3.3).

This Section begins by considering two recent models simulating the dispersal processes which result in observed distributions of fungal spores and actively moving insects. It then discusses a number of commonly used simple descriptive models of dispersion, one of which is then used to examine the consequences of spatial distributions on the dynamics of an aphid-parasitoid system.

3.2.2 Passive dispersal: fungal spores

As in Section 3.1, most models of fungal epidemics simulate the process of spore production and dispersal by means of a single rate variable: the effective number of spores per mother lesion per day. Clearly, this involves considerable simplification (in any case, often unavoidable, owing to the lack of quantitative data on the biological and physical backgrounds). However, the dispersal process itself depends on a wide range of physical factors, and is sufficiently complex to justify the use of numerical simulation methods.

This section considers such a more detailed model of the dispersal process alone. It simulates the dispersal of airborne spores of the basidiomycete, *Chondrostereum purpureum* in woodland (Wagenmakers, 1984). The basidiomycete *Chondrostereum purpureum* is a common fungus in temperate regions. It occurs as a saprophyte on many deciduous trees, and as a parasite of fruit trees and various ornamentals, where it causes silver leaf disease. Spores landing on fresh wounds develop into mycelia, which obstruct the sieve tubes and, as a result, the plant produces gum in the sapwood. The fungal toxins cause senescence and release of the leaf epidermis, resulting in the characteristic silver leaf disease symptoms.

In the autumn, basidiocarps are produced (these are visible as a purple crust on the wood). These basidiocarps produce spores, which are ejected and subsequently dispersed by air currents.

Prunus serotina Erhr (American black cherry) is a tree species, introduced into Europe in the first part of this century from the United States, to improve the understory of woodlands on poor sand soils. This species, however, has become a serious pest, as it prevents native species from germinating and developing.

P. serotina is sensitive to *C. purpureum* and can be controlled by cutting the trees and treating the wounds with a suspension of mycelium or spores (de Jong & Scheepens, 1985; de Jong, 1988). Newly-formed shoots show the characteristic silver leaf disease symptoms, and die off in a few months; 60 to 70% of the uninfected branches on treated trees also die.

Before introducing *C. purpureum* as a biological control agent, however, plant protection authorities considered it necessary to estimate the increased chance of adjacent plots becoming infected. The 'normal' dispersal of *C. purpureum* should, therefore, be compared with the dispersal in areas infested as a result of control of *P. serotina*. Large-scale field trials are laborious and risky, so instead, small-scale trials were combined with theoretical calculations to evaluate the consequences of this way of control. Computer models were used to simulate spore dispersal. In the small-scale trials, results of the computer model are tested against experimental results. The validated computer model can be used to calculate infection risks at various times and under various conditions.

The model presented here is based on the micrometeorological studies of Goudriaan (1977), which simulates the dispersal of *C. purpureum* spores in a coniferous forest, but can be easily adapted to simulate other conditions.

Structure of the model To simulate the dispersal of *C. purpureum* correctly, the forest must be divided into at least four spatial sections; a spore production layer, a stem layer, a crown layer and the air above the canopy. The layers are schematically represented in Figure 28. Each layer is characterized by its own specific biological and aerodynamic properties.

Spore production takes place on the stumps that have been infected with a sporal substrate. An epidemiological model of the fungus (Section 3.1) could be used to study the epidemic on the stumps within this spore production layer.

Spore production and dispersal processes are schematically represented in Figures 28 and 29. Spore production is held proportional to the amount of sporulating area on the cut stumps: the basidiocarp area. The relation between temperature and spore production rate per unit of basidiocarp area is derived from the data of Grosclaude (1969).

Wind is responsible for the dispersal of spores to other air layers outside the forest. Spores are exchanged vertically between the layers of air by turbulent air movements.

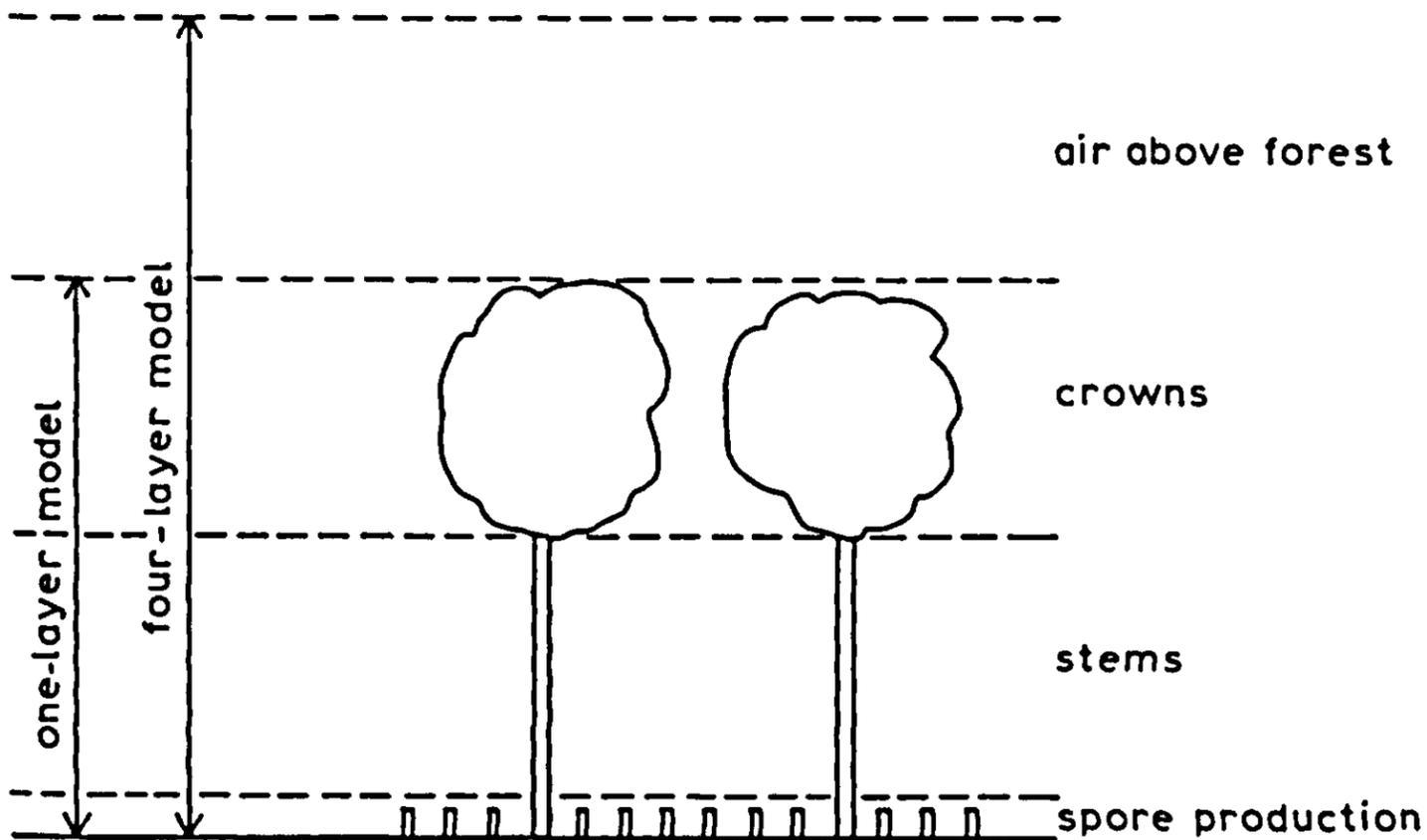


Figure 28. Layer structure of the forest.

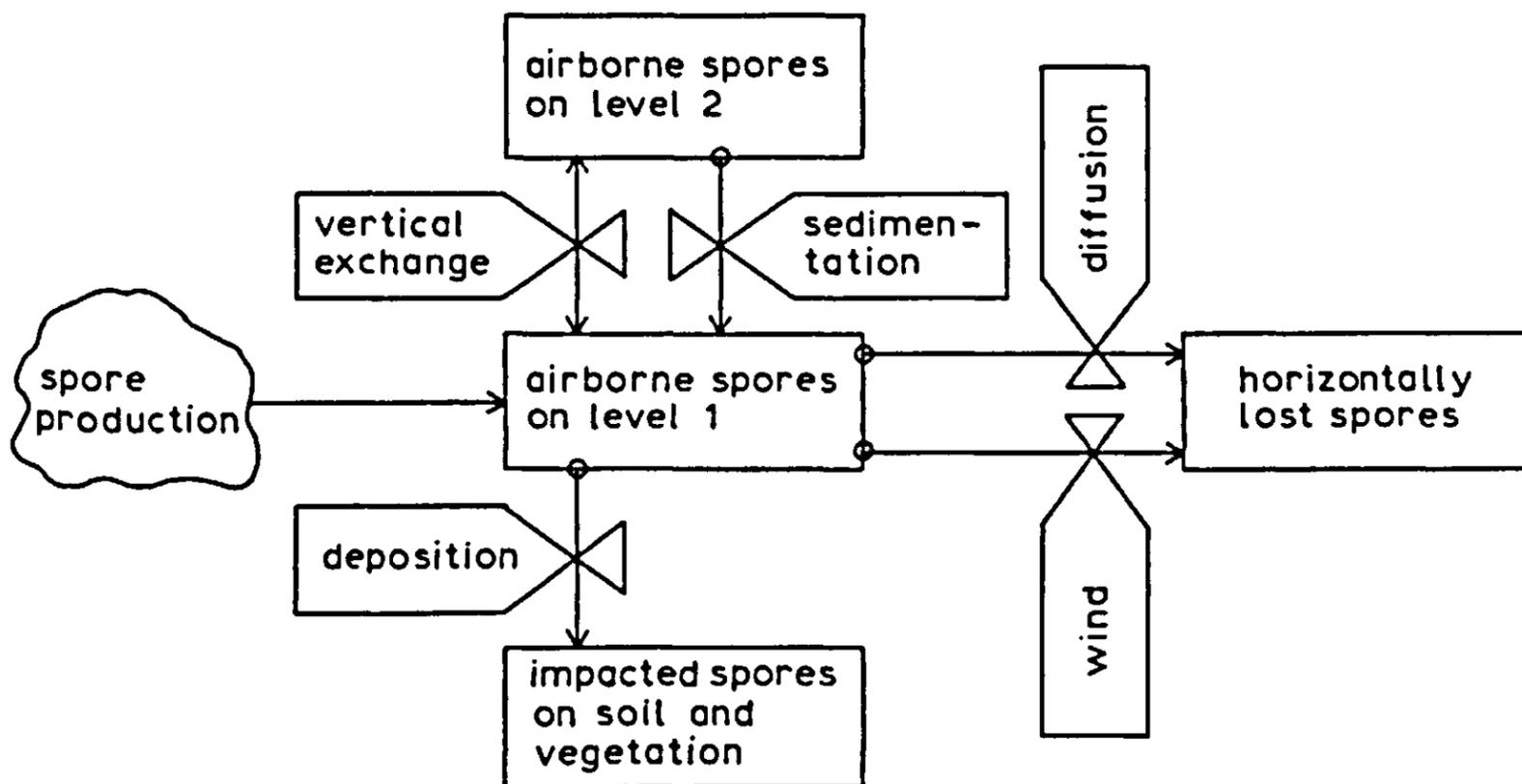


Figure 29. Relational diagram of the transport processes.

The rate of exchange is determined by exchange resistances, which have the units m^{-1} . The spore flux between two adjacent layers ($\text{spores m}^{-2} \text{s}^{-1}$) is found by dividing the difference between the spore densities by the exchange resistance.

A multi-layer model requires a lot of information on the various exchange rates, together with a detailed description of the aerodynamics of a canopy. Therefore, a simplified approach, using a one-layer model, is used in this book.

One-layer model The multi-layer model can be developed later from a simplified one-layer model in which the crown, stem and stump layers are grouped together

into one mixed layer, exchanging spores with the open air above (Figure 28). In this simplified model, the net rate of change in the number of spores inside the forest is summarized as

$$\frac{dS}{dt} = R_{\text{prod}} + R_{\text{ex}} + R_{\text{sed}} + R_{\text{dep}} \quad \text{Equation 42}$$

where S is the number of spores per m^2 of ground area, R_{prod} the rate of spore production on the stumps in numbers $\text{m}^{-2} \text{s}^{-1}$, R_{ex} the net rate of exchange by turbulent air movements in numbers $\text{m}^{-2} \text{s}^{-1}$, R_{sed} the rate of sedimentation due to settling of spores in numbers $\text{m}^{-2} \text{s}^{-1}$, and R_{dep} the rate of deposition of spores on leaves, branches etc. in numbers $\text{m}^{-2} \text{s}^{-1}$.

The concentration of spores inside the forest c_i (m^{-3}) is the number of spores per m^2 ground area divided by the thickness L of the layer (the height of the forest):

$$c_i = S/L$$

The rates R can be formulated as:

$$R_{\text{ex}} = (c_a - c_i)/r_{\text{ex}}$$

$$R_{\text{sed}} = -v_{\text{sed}} c_i$$

$$R_{\text{dep}} = -u \varepsilon_{\text{dep}} \text{LAI} c_i$$

where c_i is the spore concentration in the layer (m^{-3}), c_a the background concentration of spores in the air above the forest (m^{-3}), r_{ex} the exchange resistance between the open air and the forest (s m^{-1}), v_{sed} the sedimentation velocity (m s^{-1}), u the wind speed inside the forest (m s^{-1}), ε_{dep} the deposition efficiency of leaves (-), and LAI the leaf area index ($\text{m}^2 \text{m}^{-2}$).

Typical values for the parameters in these equations are: $R_{\text{prod}} = 20 \text{ m}^{-2} \text{ s}^{-1}$; $L = 17 \text{ m}$; $c_a = 100 \text{ m}^{-3}$; $r_{\text{ex}} = 20 \text{ s m}^{-1}$; $v_{\text{sed}} = 0.001 \text{ m s}^{-1}$; $u = 2 \text{ m s}^{-1}$; $\varepsilon_{\text{dep}} = 0.01$; and $\text{LAI} = 2$.

Equation 42 can now be rewritten in a form that groups the influx and efflux:

$$L \frac{dc_i}{dt} = \left(R_{\text{prod}} + \frac{c_a}{r_{\text{ex}}} \right) - \left(\frac{1}{r_{\text{ex}}} + v_{\text{sed}} + u \varepsilon_{\text{dep}} \text{LAI} \right) c_i \quad \text{Equation 43}$$

or, more briefly:

$$L \frac{dc_i}{dt} = b - ac_i \quad \text{Equation 44}$$

Exercise 37

Find the values of b and a by substituting the given parameters. What are their units? What is the value of the time coefficient of this equation?

At equilibrium the rate of change is zero, so spore concentration in the layer is given by:

$$c_{eq} = b/a \qquad \text{Equation 45}$$

Exercise 38

Calculate the value of c_{eq} using Equations 43, 44 and 45 with the given typical parameter values. How long does it take before c_i lies within 5% of c_{eq} , when it starts at the value c_a .

This model, despite its simplicity, gives a good impression of the relative importance of the different processes of spore dispersal. The coefficient, a , is the result of exchange, sedimentation and deposition. The sedimentation rate appears to be dwarfed by turbulent exchange and deposition. These latter two processes are of roughly equal importance, so about half the spores produced are deposited.

Above the forest, the spore concentration is set arbitrarily at $100 \text{ spores m}^{-3}$. The concentration inside the forest depends on both the concentration above the forest and on the spore production rate.

For a forest with no sporulating surface, the removal processes will reduce c_{eq} below c_a . Here, when R_{prod} is zero, so that $(R_{prod} + c_a/r_{ex})$ equals $0 + 100/20 = 5$, c_{eq} will reach a value of $5/0.091$, or only 55 spores m^{-3} . Thus, the concentration inside the forest, will be about half that in the air above.

Losses of spores downwind The model developed so far has considered an infinitely extended forest, for which c_{eq} is not a function of horizontal distance. However, at the upwind side of the forest, air blows into the forest carrying spores at the background concentration c_a (Figure 30). On a transect through the forest, in the downwind direction, spore concentration will gradually rise, until it finally reaches the equilibrium concentration c_{eq} , calculated above. The time coefficient of this process depends on the variables L and a (see Exercise 37). After about three of these time coefficients have elapsed, equilibrium will be practically established. In the meantime, the wind will have traversed a distance to be

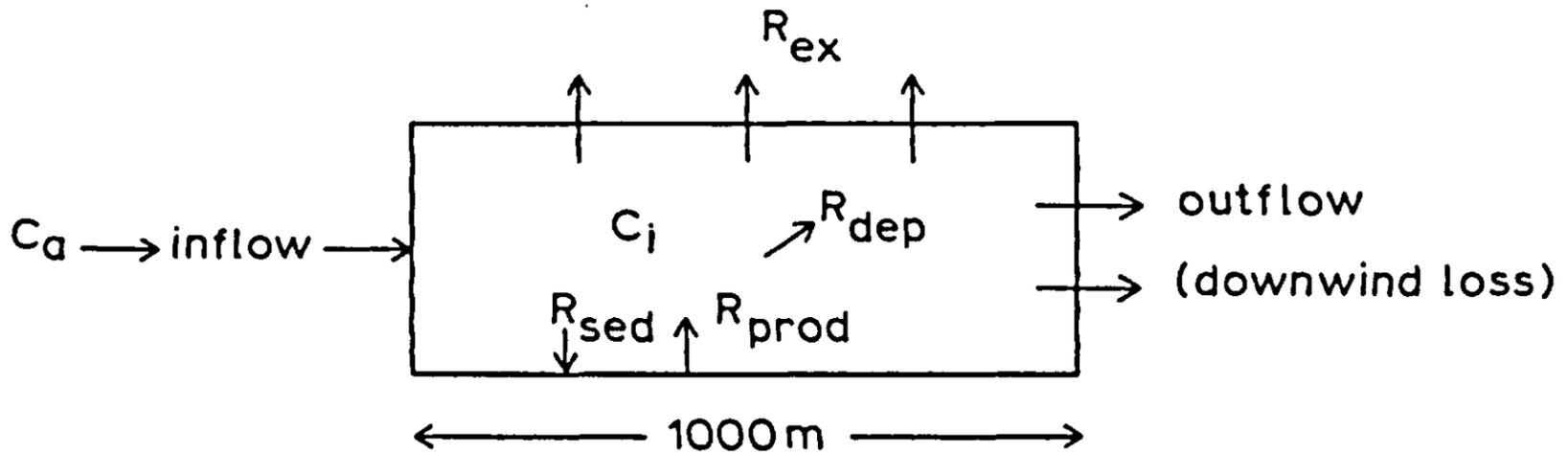


Figure 30. Rates of inflow, outflow, upward exchange (R_{ex}), sedimentation (R_{sed}), deposition (R_{dep}), and production (R_{prod}) of spores in a forest with a one-layer model.

calculated as $3 \cdot u \cdot L/a$, or about 1000 m. This means that for a forest of less than about 1000 m wide, c_{eq} is not reached.

The inflow of spores at the upwind side is about $L \cdot u \cdot c_a$, and the outflow at the leeward side about $L \cdot u \cdot c_{eq}$, so that the net downwind loss is about $L \cdot u \cdot (c_{eq} - c_a)$. This quantity can be calculated to be $5950 \text{ m}^{-1} \text{ s}^{-1}$, for the parameters chosen above. The unit m^{-1} in this expression means that the rate is expressed per m of forest width, perpendicular to the wind direction. Total spore production over a forest that is 1000 m wide is in the same units $R_{prod} \cdot 1000$ or $20000 \text{ m}^{-1} \text{ s}^{-1}$. The lateral loss thus amounts to more than a quarter of the total production. The upward loss is more difficult to calculate, since c_i actually varies over the forest width from c_a to c_{eq} . Horizontal compartmentalization would be required to find the horizontal gradients, and the total loss in a vertical direction.

This model accounts for the finite limits of a forest and is thus more appropriate for application in actual situations. Another extreme is a very small forest, i.e. a point source of spore production. This situation is not considered, as the widely used Gaussian plume model can then be applied. The one-layer model is used as a first approximation and is adequate if only temperature and humidity are to be considered, but is inadequate for spores that are produced in the spore production layer. Therefore, a four-layer model, or other more complicated model, has to be used.

Four-layer model Although the processes are essentially the same, a four-layer model (Figure 28) is complicated by spore exchange between the layers. This means that in a rate equation like Equation 43, spore concentrations must appear in adjacent layers. The system must now be represented as a set of four equations, which can be written in matrix form as:

$$\vec{L} \cdot \frac{d\vec{c}}{dt} = \vec{B} - |A| \vec{c} \quad \text{Equation 46}$$

The various elements of the matrix are derived from the geometric and agronomic characteristics of the canopy. At equilibrium, the rate of change is zero, so

$$|A| \vec{c} = \vec{B} \text{ or} \quad \text{Equation 47}$$

$$\vec{c} = |A^{-1}| \vec{B} \quad \text{Equation 48}$$

The matrix $|A|$ contains the removal processes, and the exchange resistances between the layers. The vector \vec{B} contains the source terms which are independent of the spore concentrations inside the forest. These source terms represent the rate of spore production in the bottom layer and the gross influx of spores from the air above. When the terms of A and B are substituted, Equation 47 becomes

$\frac{1}{r_1} + v_{sed}$	$-\frac{1}{r_1} - v_{sed}$	0	0	×	=	c_1	R_{prod}
$-\frac{1}{r_1}$	$\frac{1}{r_1} + \frac{1}{r_2} + v_{sed}$	$-\frac{1}{r_2} - v_{sed}$	0			c_2	0
0	$-\frac{1}{r_2}$	$\frac{1}{r_2} + \frac{1}{r_3} + v_{sed} + u_3 \epsilon_{dep} LAI$	$-\frac{1}{r_3} - v_{sed}$			c_3	0
0	0	$-\frac{1}{r_3}$	$\frac{1}{r_3} + \frac{1}{r_4} + v_{sed}$			c_4	$\left(v_{sed} + \frac{1}{r_4}\right) c_a$

The subscripts, 1 to 4, refer to the stump, stem and crown layers and an open air layer above. Typical values for the resistances r_{1-4} are $r_1 = 200$, $r_2 = 50$, $r_3 = 10$, $r_4 = 4 \text{ s m}^{-1}$, respectively; u_3 will be about 2 m s^{-1} . The other parameters are the same as before.

Exercise 39

Substitute these parameters (r_{1-4} , u_3) into Equation 47 and set up the numerical form of the equation. Determine the vector c , preferably using an existing software package.

The resulting vector c , which can be calculated by matrix inversion from Equation 48, gives the spore concentrations in the four layers. As shown by Wagenmakers (1984), these concentrations differ by factors of 10. The one-layer model is, therefore, a gross oversimplification.

Influence of weather and forest structure The exchange resistances are approximately inversely proportional to windspeed. Since v_{sed} is small, the entire matrix A is almost proportional to windspeed, so c_{eq} will be roughly inversely proportional to windspeed. Temperature and humidity affect the system, mainly as a result of their influence on the rate of spore production. Other parameters being constant, $(c_{eq} - c_a)$ will be proportional to R_{prod} .

In a denser forest, the intercepting area (LAI) will be larger, but windspeed in the crown layer will fall. The product $u_3 \cdot LAI$ may change either way. The exchange resistances will be larger, so a steeper spore gradient will develop. An extensive sensitivity analysis is given by Wagenmakers (1984).

3.2.3 *Active dispersal in two dimensions*

Unlike fungal spores, animals move actively. Therefore, spatial dispersal of these organisms should be considered in a different way. In predator-prey systems, the movement of prey and predator is decisive for the number of encounters between them, and this determines the predation rate of individual predators. Acarine systems are well-studied predator-prey systems. They are found in various crops and, in many cases, enable feasible biological control. To understand such systems, the individual predator's dispersal is studied. It appears that the movement of an individual predatory mite is determined by its condition, often expressed in gut content. The more a predatory mite has eaten, the less its linear displacement. This is explained in a simulation study of predatory mite movement by Sabelis (1981), based on a detailed analysis of walking patterns in relation to gut content.

The model presented in this Subsection simulates the walking patterns of predatory mites searching for spider mites, *Tetranychus urticae*, on rose leaves. The formulation, however, can be used for a wide range of organisms. The resulting simulation models can be used as a basis for determining predation rates of individual predators, when predation rate is mainly determined by encountering rate and success ratio. These individual predation rates may be used in models of pest population dynamics. The mite's continuous movement is modelled as a series of linear steps of constant length. This allows separation of the two determinants of the walking pattern: speed and direction. Here, we will consider only the directional component.

Figure 31 illustrates the division of the mite's curved path into linear steps of fixed length. The angular deviation between the directions of steps $s-1$ and s is denoted by A_s (in radians). It is important to note that A_s expresses the change in direction, rather than the direction itself.

The angular deviation per step can be used as follows, to generate a two-dimensional walking pattern. The direction (a state variable) is calculated in a CSMP program as (see Appendix 5):

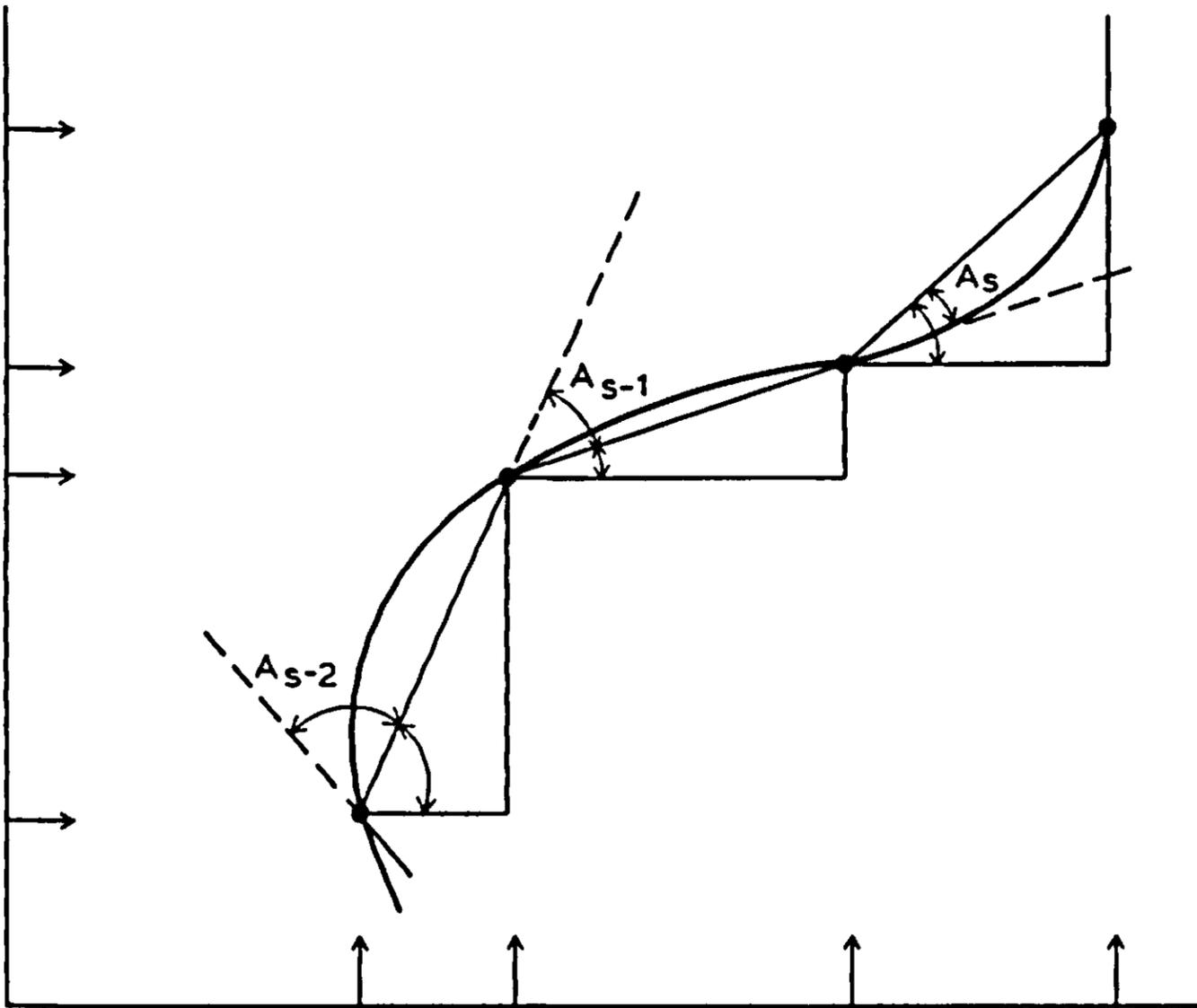


Figure 31. Determination of three consecutive angular deviations (A_{s-2} , A_{s-1} , A_s) after pacing linear distances of fixed length. (Source: Sabelis, 1981).

$$\text{DIR} = \text{INTGRL}(\text{DIRI}, A/\text{DELTA}) \text{ or mathematically } \text{DIR} = \int (A/\Delta t) dt$$

where DIRI is the initial direction in radians, relative to a reference direction, A is the angular deviation (A_s) and DELTA is the time step of integration (DELTA).

To calculate movement in two dimensions, we write in CSMP:

$$\text{CSDR} = \text{COS}(\text{DIR})$$

$$\text{SNDR} = \text{SIN}(\text{DIR})$$

$$X = \text{INTGRL}(X_0, \text{CSDR}/\text{UNIT}) \text{ or mathematically } X_t = \int_0^t (\text{CSDR}/\text{UNIT}) dt$$

$$Y = \text{INTGRL}(Y_0, \text{SNDR}/\text{UNIT}) \text{ or mathematically } Y_t = \int_0^t (\text{SNDR}/\text{UNIT}) dt$$

The term UNIT is required because the magnitudes of CSDR and SNDR cannot exceed unity, whereas X and Y are the distances from the start of the step. The time step DELTA dictates the physical time step taken by a predator, so that this

time step of integration should be based on the step size of the predator. UNIT thus determines the organism's speed of movement.

We can now consider the simulation of the successive values of A_s , the angular deviation. First, it should be noted that A_s is the angular deviation per step. This means that the measurements used in constructing the model cannot be analysed until the step length has been decided. The modelling approach is thus required early on in the study; it is not sufficient to collect data first and make the model later.

The rate of turning is described in terms of two components, one of which depends on previous movements, the other being considered as a random 'error term'. A_s can be expressed as the dependent variable of a multiple auto-regression:

$$A_s = a_1 \cdot A_{s-1} + a_2 \cdot A_{s-2} + \dots + a_m \cdot A_{s-m} + E_s$$

where a_1, a_2, \dots, a_m are the regression coefficients, and E_s is the random 'error term'; m is called the order of the process, and is a measure of the length of the animal's 'memory'. These higher order terms account for the memory of the animal that affects its present step direction.

Exercise 40

Give the equation for a first-order walking process. What is the effect of changing the sign of the coefficient from negative to positive?

In the simulation, E_s may be drawn at random from a specified distribution. The Tukey distribution is most appropriate since, by changing its three parameters, it can be made to approximate several other well-known distributions (Sabelis, 1981). This distribution can be generated as follows:

$$E_s = \mu + \sigma \cdot Y_p$$

$$Y_p = \begin{cases} (p^\lambda - (1-p)^\lambda)/\lambda & \lambda \neq 0 \\ \ln(p/(1-p)) & \lambda = 0 \end{cases}$$

where μ is a parameter determining the mean change in direction (independent of previous turns), σ determines the variance in the angular deviations, and λ is a parameter related to the distribution's kurtosis. (A distribution with high kurtosis has a relatively sharp peak and long flat tails.) The variable p is the cumulative frequency, and is drawn at random from a uniform distribution ($0 \leq p \leq 1$). The approximated distribution depends on the value of λ (Snedecor & Cochran, 1980), as follows:

$\lambda = 1$	uniform
$\lambda = 0.14$	normal
$\lambda = 0$	logistic
$\lambda = -0.85$	Cauchy

Exercise 41

How well does the Tukey approximation compare with a uniform distribution? What are the maximum and minimum values of E_s ? What behaviour is expressed by μ ?

The angular deviation in a first-order walking process is now given by

$$A_s = a \cdot A_{s-1} + \mu + \sigma \cdot (p^\lambda - (1-p)^\lambda) / \lambda$$

In fact, many observed walking patterns are 'zero order' (i.e. the change in direction during step s is independent of previous turns) so the coefficient a equals 0. A_s can thus be calculated as follows in CSMP:

$$P = \text{RNDGEN}(U)$$

$$A = \text{MU} + \text{SIGMA} * (P ** \text{LAMBDA} - (1. - P) ** \text{LAMBDA}) / \text{LAMBDA}$$

where $\text{MU} = \mu$, $\text{SIGMA} = \sigma$, $\text{LAMBDA} = \lambda$; $U =$ an odd integer, so that $0 \leq P \leq 1$.

Figure 32 presents some examples of the walking patterns generated by this model, with different values of σ and λ . These figures represent the patterns of single individuals, although they were generated using parameters which describe the range of behaviours measured in a population (e.g. μ , σ , the regression coefficients a_i).

The use of such a model to determine the effects of an individual's behaviour on changes at the population level (e.g., dispersion, effects on prey density etc.), involves two main steps. First, the consequences of an individual's action must be calculated; then these effects must be summed or averaged over the whole population.

Clearly, an individual's behaviour has a wide range of consequences: dispersal, rate of encounter with prey individuals (which also depends on the distribution of the prey population) or mates, etc. The consequences to be modelled thus depend on the purpose of the model.

3.2.4 Descriptive models and the consequences of dispersion

The previous two Subsections have considered techniques for modelling dispersal – the process by which a population's dispersion pattern develops. In many cases, however, quantitative data at the process level are insufficient to

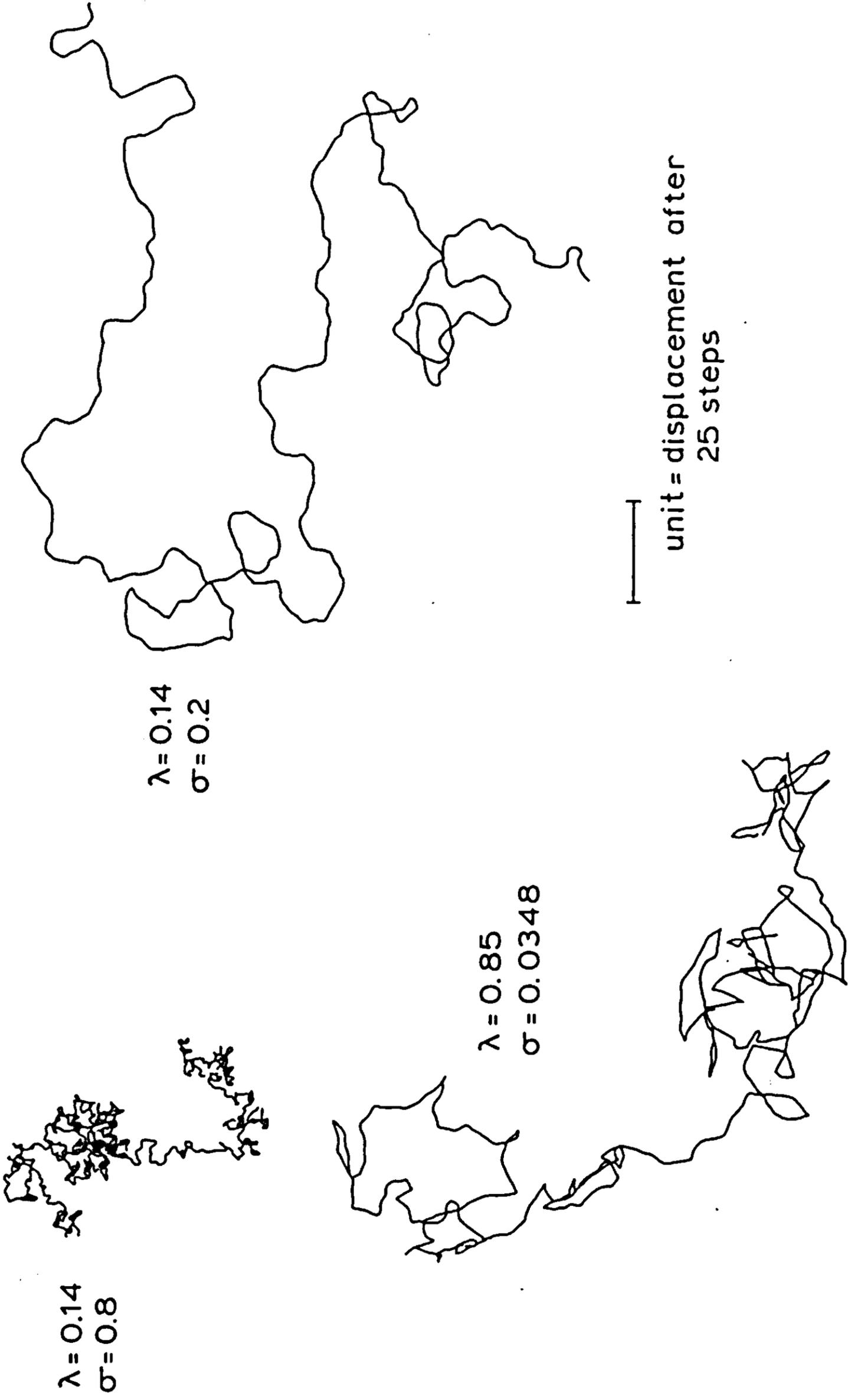


Figure 32. Some walking pattern matched by simulation. (Source: Sabelis, 1981).

construct an explanatory model of the resulting behaviour and dispersion. It is also usually unnecessary to use detailed explanatory models of individuals behaviour to predict effects at the population level: simpler descriptive models of the distribution of individuals in space are often sufficient as a basis for models of population dynamics (see Section 3.3). Although these descriptive functions may be summaries of the results of more detailed explanatory simulation models, the lack of appropriate behavioural measurements often means that simple statistical distributions must be used as a basis for studying the consequences of dispersion. Such consequences may concern the possibilities of biological control in various distributed prey populations. This is illustrated in Subsection 3.2.5.

Before presenting some of the more commonly used models, three cases should be distinguished:

1. A population of distinct individuals, distributed among discrete sample units (e.g. aphids distributed among wheat tillers);
2. Distinct individuals distributed over a continuous area or volume (e.g. mites on a leaf, nematodes in soil, or weed seeds in a field);
3. Indistinct 'individuals' distributed over a continuous area (e.g. fungal lesions on leaves).

In models simulating the consequences of dispersion, cases (2) and (3) are often considered as if they were populations of type (1); continuous areas are divided into 'discrete' squares. The reason for this is that most distribution models can be applied only to case (1).

We can now consider some of the statistical distributions commonly used to describe the spatial distributions of pest and disease populations.

Spatial distributions Considerable information on spatial distribution of individuals is given in most ecological textbooks; e.g. Southwood (1978). Here, only the most common distributions will be discussed.

The Poisson distribution This arises if organisms are distributed at random i.e., the probability of an individual being 'placed' in a particular sample unit is independent of the number of individuals already there, and does not vary among sample units. The probability that a sample unit contains x individuals is given by:

$$P_x = \mu^x \cdot e^{-\mu} / x!$$

where μ is the mean density per unit. An important feature of this distribution is that the between-unit variance in density is equal to the mean.

Generalized distributions These can arise if individuals are clustered. The most commonly used such distribution is the Negative Binomial Distribution (NBD), in which cluster size follows a logarithmic distribution, and clusters are distrib-

uted at random (Poisson). Here, the probability that a cluster contains i individuals is

$$P_i = -a^i / (i \cdot \ln(1 - a)) \quad \text{Equation 49}$$

where a is a constant ($0 < a < 1$). The number of clusters per sample unit is a Poisson variable.

The main features of the NBD are, in terms of its two parameters μ (mean density) and k (which describes the degree of clustering),

$$V = \mu \cdot (1 + \mu/k) \quad \text{Equation 50}$$

$$P_x = \frac{(k + x - 1)! \cdot \mu^x}{k^x \cdot (k - 1)! \cdot x!} \cdot (1 + \mu/k)^{-(k+x)}$$

where V is the between-sample unit variance in the number of individuals, and P_x the probability that a sample unit contains x individuals.

The full NBD can be generated from:

$$P_0 = (1 + \mu/k)^{-k} \quad \text{Equation 51}$$

$$P_{x+1} = \frac{P_x \cdot \mu \cdot (k + x)}{k \cdot (x + 1) \cdot (1 + \mu/k)} \quad \text{Equation 52}$$

For the derivation of Equations 51 and 52 the reader should refer to the ecology textbooks already mentioned.

Fitting a distribution to observed data Clearly, before testing whether field data can be described by one of these distributions, it is necessary to estimate the parameters. Estimation of the mean, μ , from the sample mean (m), e.g. the number of individuals per tiller, per leaf, m^{-2} , or per plant, is sufficient to generate a Poisson distribution; but the NBD requires additional parameters. Besides estimating the mean, the k -value characterizing the level of clustering should also be estimated.

Three methods are commonly used to estimate k for an observed distribution (Southwood, 1978). The first uses iteration to solve the equation (see Equation 51):

$$\log P_0 = -k \cdot \log(1 + m/k) \quad \text{Equation 53}$$

where m is the sample mean (an estimate of μ) and P_0 is the proportion of unoccupied sampling units (e.g. wheat tillers carrying no aphids). The second method uses the relation (see Equation 50):

$$k = m^2 / (V - m), \quad \text{Equation 54}$$

where V is the between-tiller variance in aphid density. Finally, a maximum likelihood estimate of k can be found by iterative solution of the equation

$$N \cdot \ln(1 + m/k) = \sum_x \left(\frac{A_x}{k + x} \right) \quad \text{Equation 55}$$

where N is the number of sample units, and A_x the number of sample units carrying more than x individuals (e.g. $A_6 = f_7 + f_8 + f_9 + \dots$). Equations 53–55 are derived from the ecological textbooks.

Having estimated the appropriate parameter values, the measured data can be compared with the predicted distribution. The simplest way of showing to what extent the data fit the predictions, involves comparing the moments (variance, skewness and kurtosis) of the observed distribution with those of the predicted form.

If the data conform to a Poisson distribution, the sample variance will be equal to the sample mean, i.e. $V/m = 1$. This is easily tested, since $N \cdot V/m$ is distributed approximately as a χ^2 variate with $N - 1$ degrees of freedom (N is the number of sample units).

To test whether observed data conform to an NBD with known k , either the sample variance or the skewness can be used, depending on the values of m and k . Details of these methods are again given in Southwood (1978).

Exercise 42

Table 8 gives data on the frequency of *Drosophila melanogaster* (Meigen) containing various numbers of eggs of the parasitoid *Pseudeucoila bochei* (Weld) from an experiment described by van Lenteren et al., 1978. Do these data support the hypothesis that all hosts, whether previously infested with parasites or not, are equally attractive to the parasitoids?

Table 8. Distribution of eggs of the parasite *P. bochei* among larvae of *D. melanogaster*. (Source: van Lenteren et al., 1978).

Number of egg per host larvae	Frequency
0	0
1	8
2	19
3	19
4	4
5+	0

The comparison of observed and predicted distributions can only be used to test the descriptive accuracy of a particular model. If the data differ significantly from the predictions, then at least one of the model's assumptions must be untrue. A close correspondence between observations and predictions, however, does not confirm the assumptions underlying the theoretical distribution, and cannot be used as a basis for inferring that the mechanisms result in an observed distribution. For example, a Negative Binomial Distribution may result from the random distribution of clusters (whose size is distributed according to Equation 49) among identical sample units, or from the random distribution of individuals among sample units of varying suitability (see Pielou, 1977).

3.2.5 *Effect of dispersion on the dynamics of an aphid-parasite system, an example*

The system to be considered here involves the cereal aphid, *Sitobion avenae* on winter wheat, and its hymenopterous parasitoid *Aphidius rhopalosiphi*. The original model of population dynamics assumed implicitly that the aphids and parasites were distributed uniformly among wheat tillers. To test the effects of introducing more realistic assumptions about the spatial distributions of the two species, a simulation model that takes the measured distributions into account was developed (Rabbinge et al., 1984a).

A series of steps is involved:

1. Selecting the appropriate statistical distribution, for describing the populations' dispersion;
2. Calculating the distribution parameters;
3. Determining a method for predicting these parameters;
4. Generating the populations' distribution; and
5. Calculating the resulting parasitism rate.

Choice of a distribution (1) Field data on *S. avenae*, when compared with the results predicted by the Poisson, Negative Binomial, and Neyman A distributions, have been shown in most cases to be best described by the Negative Binomial Distribution (Rabbinge & Mantel, 1981). The model, therefore, assumes that this is the appropriate form.

The parasites are assumed to be dispersed according to one of two extreme distributions: either uniformly (i.e., the likelihood that a parasite is present on a tiller is independent of the number of aphids present) or in proportion to the number of aphids on each tiller.

Estimation of the parameters (2) As stated above (Subsection 3.2.4), the Negative Binomial Distribution is characterized by two parameters: the mean, μ , and k , which determines the degree of aggregation of the population. The mean density will be generated by the population model, so need not be estimated here. The parameter, k , however, must be determined on the basis of field observations.

Modelling the changes in k (3) Three methods have been used in various studies. The first relies on a descriptive relation between m and V , of the form:

$$V = a \cdot m^b \quad \text{Equation 56}$$

where a and b are constants to be estimated by regression from a series of observations on cereal aphid distribution (Taylor, 1961). This has been shown to describe the mean-variance relation for a wide range of species (Taylor et al., 1978). One parameter k can then be calculated from m (by substituting Equation 56 into 54)

$$k = m^2 / (am^2 - m) = m / (am^{b-1} - 1)$$

Alternatively, k can be estimated using Equation 53, by substituting a descriptive relation between m and P_0 (e.g. those found by Rabbinge & Mantel, 1981; Nachman, 1981).

The simplest method, which will be used here, relies on a regression of the maximum likelihood estimates of k on the mean density. Using sample data on cereal aphids, Rabbinge et al. (1984a) calculated the regression:

$$k = 0.3128 + 0.0724 \cdot m \quad \text{Equation 57}$$

Generation of the dispersion of the aphid population (4) The dispersion of the aphid population is generated in this example from the theoretically formulated curve for a Negative Binomial Distribution, which is characterized by the mean and the clustering parameter k . Within the theoretically defined distribution of individuals, various density classes are distinguished, within which the assumption of linearity for parasite-infestation rate is valid.

In this example, five density classes are considered: 0, 1–5, 6–25, 26–125, and > 126 aphids tiller⁻¹, respectively. The proportion of tillers in each class is calculated with Equations 51 and 57, using FORTRAN DO loops in a NOSORT section of the program. The mean aphid density is used first, to calculate the dispersion parameter, k . The frequencies in the various density classes are then generated.

```

K = 0.313 + 0.072* MEAN
P0 = (1. + MEAN/K)** (-K)
SUM1 = 0
P = P0
DO 1 I = 1,5
  Q = P* MEAN*(K + I - 1.)/(I*(K + MEAN))
  SUM1 = SUM1 + Q
  P = Q
1 CONTINUE

```

Here, P_0 is the frequency of tillers bearing no aphids; the final value of SUM1 is the sum of the frequencies in classes 1 to 5. Similar methods are used to calculate the frequencies of tillers in the other three groups of density classes.

The parasites can then be allotted to the various groups of tillers, either uniformly or in proportion to aphid density. Generation of the spatial distributions requires considerable computation time. Fortunately, the form of the distribution changes only gradually (Kroon & Driessen, 1982); this means that the aphids and parasites need not be redistributed each DELT (in this case, $DEL T = 0.01$ day). Instead, the dispersion section of the program is used only twice per simulated day. (This approach to the problem of 'Stiff equations' will be considered in Subsection 3.4.6).

Calculating the resulting rate of parasitism (5) Having modelled the distribution of aphids and parasites at the start of a time step, the total rate of parasitism can be calculated for the whole population. The parasitism rate (per parasitoid per day) increases with the number of aphids present on a tiller. This is modelled here using an AFGEN statement to describe the experimental results of Shirota et al. (1983) (Figure 33). The processes which determine this 'functional response' will be considered in Section 3.3. Its main features are that at low prey densities, parasitism is limited by the number of hosts available, while at higher densities it depends on the maximum rate at which a parasitoid can oviposit. The mean rate of oviposition is calculated for each density class, and averaged over all tillers (the weighting of individual rates depends on the proportion of tillers in the various density classes). The resulting mean rate of parasitism is then used to calculate the rate of transition from 'living' aphids to 'parasitized' aphids, which then forms a component of the mortality rate.

Figure 34 shows the results of three models: (1) aphids and parasites both uniformly distributed, (2) aphids distributed according to a Negative Binomial – parasites distributed uniformly, and (3) aphids distributed according to a Negative Binomial – parasites distributed in proportion to local aphid densities; i.e. most parasitoids are on the tillers with highest aphid densities.

Clearly, aggregation of parasitoids in areas of high aphid density increases the rate of parasitism, and thus reduces the peak aphid population. The magnitude of the effect, however, is small. In the conditions simulated here, therefore, increasing the parasitoids' searching efficiency does not have a significant influence on the growth of the aphid population; parasitism is thus unlikely to be a useful natural control method unless the density of parasites can be increased considerably.

3.2.6 *Discussion and conclusions*

The models introduced in this Section were chosen to illustrate approaches to modelling some of the aspects of dispersal and dispersion. Clearly, the range of phenomena covered by the term 'spatial heterogeneity' is far too great to review

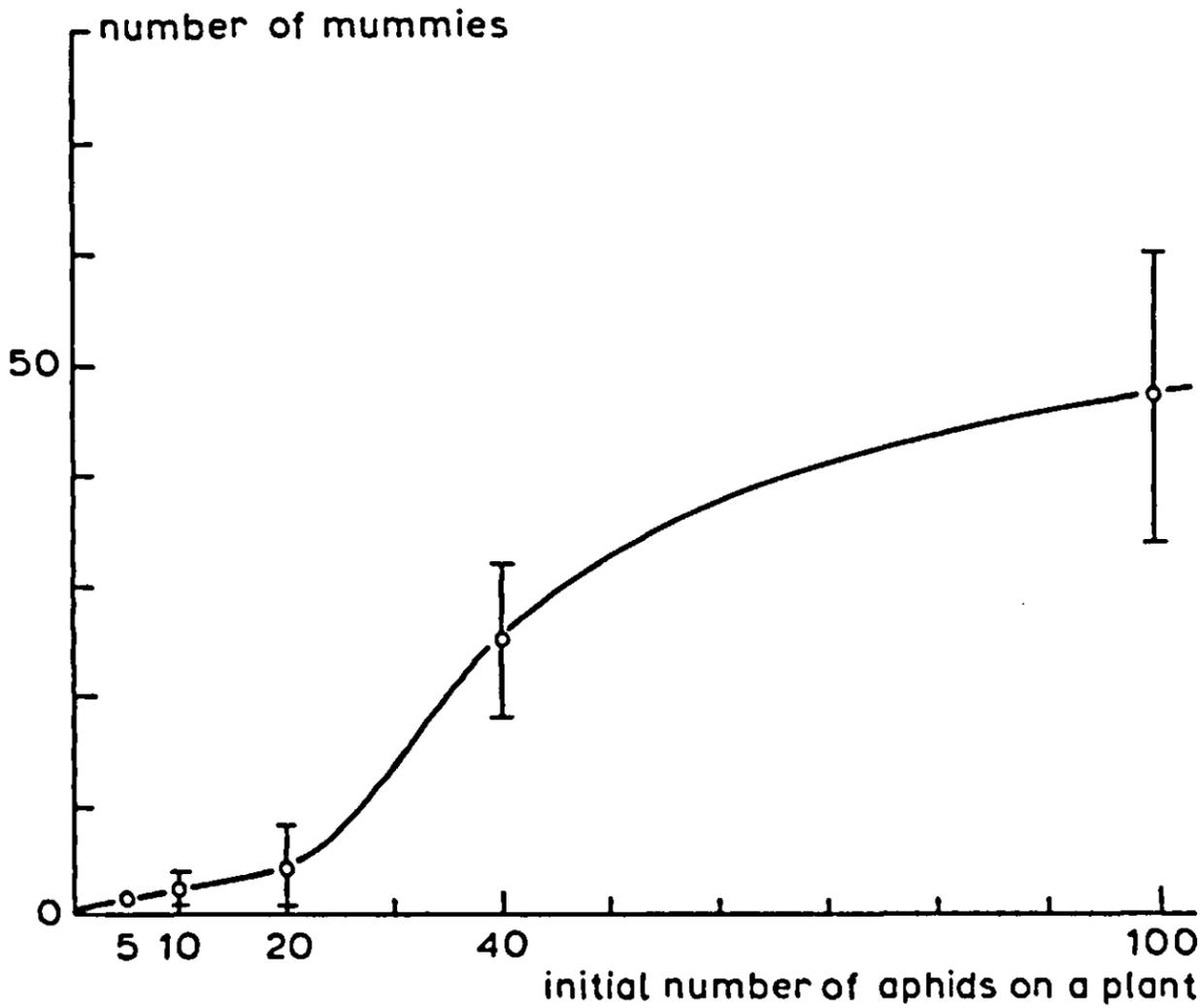


Figure 33. The functional response of *Aphidius rhopalosiphi* to changes in host density. (Source: Shirota et al., 1983).

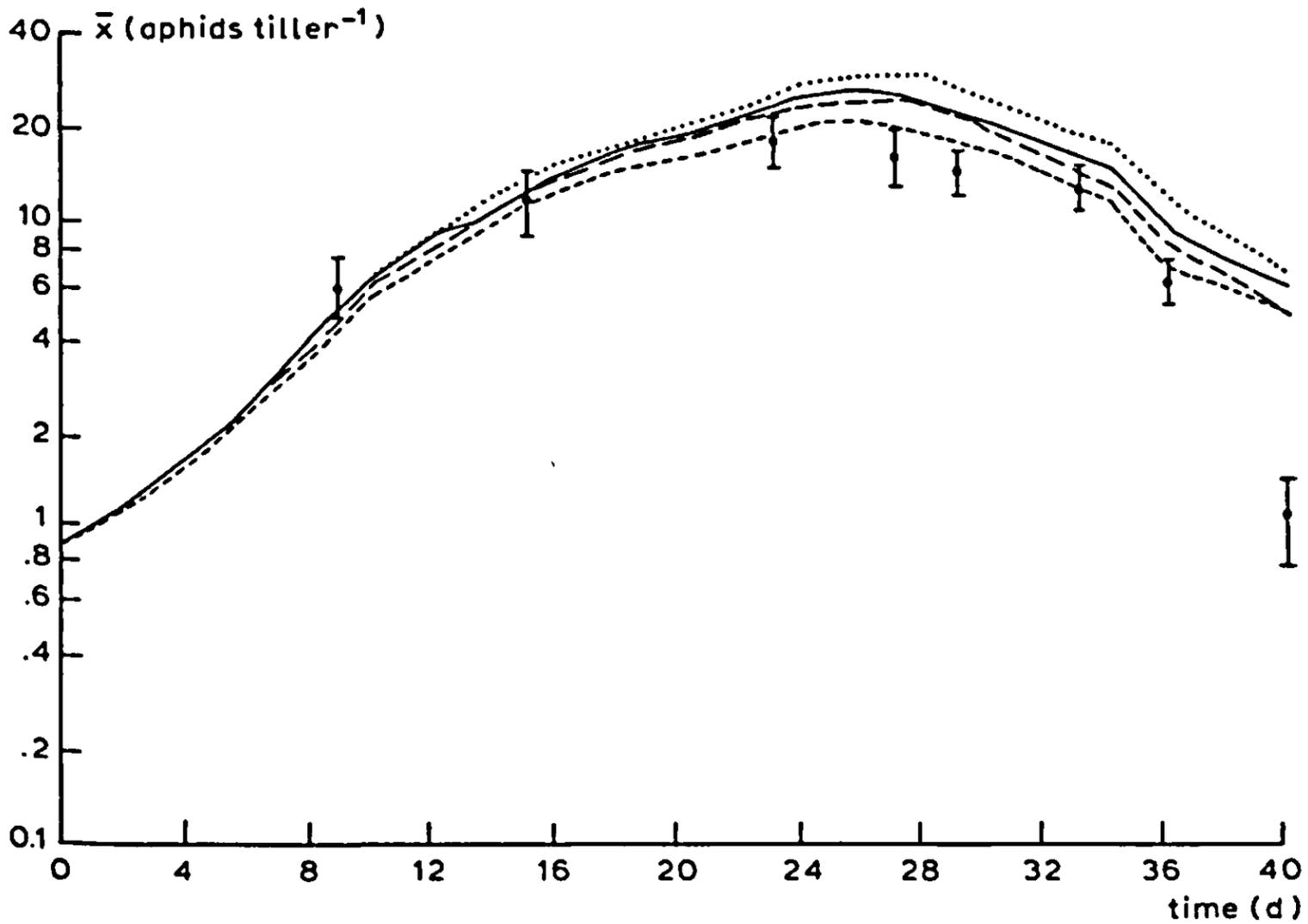


Figure 34. Mean aphid density (\bar{x}) in the field (I bar with 95% confidence intervals), and the results of simulations 1 (—), 2 (---) and 3 (-.-) (see text) and without parasitoids (....). (Source: Rabbinge et al., 1984a).

thoroughly in a single Section, but the models chosen certainly illustrate a number of important groups of features.

First, it is important to decide (in advance) whether the model is to be used to simulate dispersal (the process by which spatial distribution is produced) or the consequences of dispersion (the pattern resulting from dispersal).

Secondly, dispersal can be simulated at the level of either the individual or the population. Thus, Wagenmakers' model (1984) for passive dispersal of fungal spores simulates the process of dispersal in a system described in terms of features of the spore population; while the model for active dispersal simulates the behaviour of individual mites (Sabelis, 1981).

The other main aspect of modelling introduced in this Section, was that of evaluation of the consequences of a particular spatial distribution on the population dynamics of an organism. This was done using the model described in Subsection 3.2.5. Here, compound simulation (see also Section 3.3) was used to overcome problems of Monte Carlo analysis. The latter is needed in case stochastic processes are introduced in the models. This was not done in this Section but will be discussed in the next Section 3.3.

The very different levels and approaches of the examples considered in this Section, show that there is no standard 'right way' of simulating dispersal and dispersion. As in most of the other areas treated in this book, the choice of methods must depend on the aims of the study, the biological or physical processes concerned, and the information available.