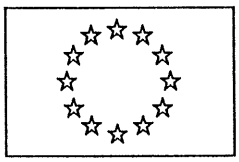

European Commission
Directorate-General XII, Science,
Research and Development
Environment Research Programme

COST 819

BIOTECHNOLOGY

Ecology and transmission strategies of entomopathogenic nematodes

Edited by
C.T. Griffin,
R.L. Gwynn,
J.P. Masson

1995 EUR 16269 EN

ANALYSIS OF SPATIAL VARIABILITY IN PEST MANAGEMENT

Wopke van der Werf[1], Tom Been[2] & Corné G. Kocks[3]

[1]Wageningen Agricultural University, Department of Theoretical Production Ecology, P.O. Box 430, 6700 AK Wageningen, The Netherlands; [2]Research Institute of Plant Protection, IPO-DLO, P.O. Box 9060, 6700 GW Wageningen, The Netherlands; [3]Wageningen Agricultural University, Department of Phytopathology, P.O. Box 8025, 6700 EE Wageningen, The Netherlands

SUMMARY

Spatially implicit and explicit techniques are used for describing aspects of spatial variability of pests, natural enemies, diseases, plants (weeds) and other biotic and abiotic factors in agricultural fields and natural terrain. The choice of technique depends upon the purpose of the study and the possibility to gather the required data. We present basic concepts underlying geostatistical analysis of disease patterns and the description of sampling and monitoring processes with probability distributions and other models. The mathematical methods are used for designing efficient pest sampling protocols. This is illustrated with case studies on the spatial pattern of a bacterial disease in cabbages, the design of monitoring systems for pest mites in apple and the detection of cyst nematode patches in potato.

INTRODUCTION

Densities of pests and their natural enemies (e.g. entomopathogenic nematodes) vary over space. Insight in the spatial dimension of pest attack is required when developing reliable and efficient methods for detecting pest presence and determining the average density. Spatial patterns may suggest the mechanisms underlying the introduction and spread of disease or pest in a field. Control measures may be targeted to hot spots, where density is highest.

Observations on the density (or some other expression of 'presence') of a pest or disease can be collected in two fundamentally different ways; *with* or *without* the spatial coordinates of the observation. In the first case, a map can be drawn of pest density in space and techniques may be used to analyse and describe the spatial pattern. In the second case, the result of the observations is a collection of numbers. Analysis and description then focus on the frequency distribution of these numbers, disregarding the spatial coordinates. When spatial coordinates are not recorded, the data become spatially implicit, i.e. spatial relationships cannot be retrieved from the

data set, although they still underly its statistical attributes. When spatial coordinates are retained, the resulting spatial analysis and relationships are *explicit.* Spatially implicit and explicit techniques are both used in research on pest ecology, but for different purposes. Spatially explicit techniques are primarily used for describing and mapping disease patterns and studying mechanisms underlying the initiation and spread of disease in field crops. A practical application of such studies is the derivation of optimal sampling distances and patterns. Spatially implicit techniques are used for describing, analysing and predicting the statistical properties of sampling methods and for developing sampling methods that strike an optimal balance between sampling effort and sampling accuracy.

The next section of this paper is methodological. We present here some important concepts in the analysis of spatial variability in pest management. The application of these concepts is illustrated in a section with three case-studies. The overall aim of this presentation is to highlight approaches that are potentially useful and can be easily adapted for the study of entomopathogenic nematodes.

METHODOLOGY

Principles of geostatistical analysis

Two approaches to *explicit* spatial analysis of pest and disease patterns dominate the phytopathological literature. One is based on geostatistics (Burrough, 1987) while the other is based on auto-regressive integrated moving average (ARIMA) models (Hudelson *et al.,* 1989). Geostatistics was originally developed for spatial interpolation and mapping in geology and mining. It is now widely used in soil science. Spatial autocorrelation analysis evolved from time series analysis. Despite their different origin, terminology and calculation methods, the two approaches have several conceptual similarities. Geostatistics is becoming an accepted technique for mapping disease patterns (Chellemi *et al.,* 1988; Lannou & Savary, 1991; Munkvold *et al.,* 1993; Nelson et al., 1994; Stein *et al.,* 1994).

The purpose of geostatistics is to create a (contour)map of the spatial pattern of a spatially varying characteristic, using interpolation between observed data points. The interpolation is done in such a way that the obtained estimates are unbiased and have minimum variance. The first step in a geostatistical analysis is the description of the statistical relationship between data points as a function of their distance. (As a rule, the correlation diminishes with distance.) The statistical descriptor of (un)relatedness, used in geostatistics, is the semi-variance:

$$\gamma(h) = \frac{1}{2} \mathrm{E}[Z(x+h)-Z(x)]^2$$

where

$\gamma(h)$ is the semi-variance for a spatial distance h

$Z(x)$ is the value for a characteristic (say pest density) at a location x

E denotes the statistical expectation

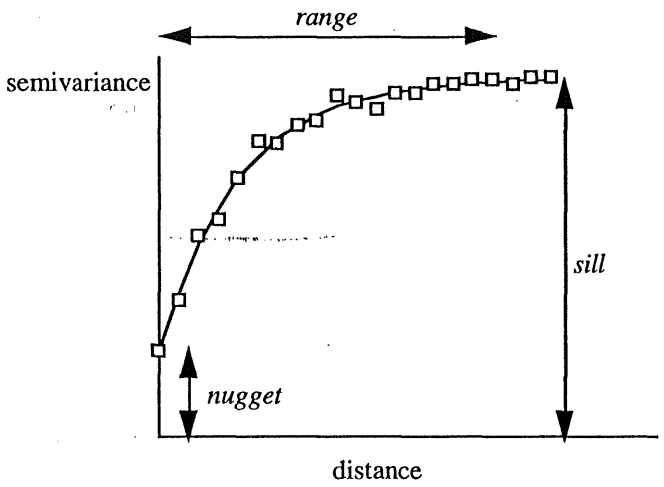

Fig. 1 Semivariogram. Horizontal axis denotes spatial distance, vertical axis the semi-variance, which is a measure of the average squared difference between observations made at a given distance. Points are calculated from a spatially indexed data set. The drawn line is a non-linear regression equation.

In a data-set of observations, collected at N different sites in a field, there are theoretically $N(N-1)/2$ estimates of the semi-variance. The data are grouped in distance classes and the semi-variance for a distance class is plotted against the distance (Fig. 1). The resulting figure is usually a curve that starts at a non-zero value for distance 0 (the *nugget*) and increases in a nonlinear way to a maximum value (the *sill*). The *range* is a measure for the distance over which the semi-variance is (substantially) smaller than the sill. A smooth curve is drawn through the data points, using nonlinear regression with an appropriate function (Fig. 1). A variety of functions describing semivariograms is used. One of those is the negative exponential:

$$\gamma(h) = nugget + (sill - nugget)(1 - \exp(-h/range))$$

When there is no spatial interdependence, the semivariogram becomes a horizontal line. This is called the *pure nugget effect.*

There are some requirements when calculating semivariograms. First, there should be at least 50 points per distance class. Second, not all the $N(N-1)/2$ data pairs may be used for the calculation of the semi-variance because the largest distance appearing in the semi-variogram should not exceed approximately half of the length of

the field. Otherwise only extreme parts of the field would be involved in the calculation, so that the result can not be regarded as representative for the whole. When constructing semivariograms, it is assumed that the semi-variance is a function of distance only, and that the variance is constant over space. Violations of these assumptions can be solved by a.o. transformation of data, using moving averages, or representing large scale variation in the underlying mean by fitting a spatial response surface (Burrough, 1987).

Spatial interpolation between observation points is done with a technique called *kriging* after one of its developers, the South African mining engineer D.G. Krige. Kriging estimates are a linear sum of weighted observations within a certain neighbourhood:

$$\hat{Z}(x_0) = \sum_i w_i Z(x_i)$$

where $\hat{Z}(x_0)$ is the interpolated function value at location x_0 and w_i is the weight of the ith measurement $Z(x_i)$ The weights depend on the semivariogram. They can be positive or negative and their sum is 1. They are determined such that the kriging estimate of $\hat{Z}(x_0)$ is unbiased and has minimum variance. The actual accuracy of $\hat{Z}(x_0)$ depends on the shape of the semivariogram, and on the density and spatial pattern of the observations.

With use of kriging; it is possible to predict values at unvisited locations. This prediction is based on neighbouring observations and the configuration of these observations. Kriging finally yields a (contour)map of estimated values and their associated prediction errors. More detailed information is given by Journel & Huijbregts (1978).

Principles of sampling and monitoring

Interest in pest sampling was spawned by the concept of Integrated Pest Management, which emerged in the Western world in the 1960s as a reaction to the alarming side-effects of chemical pest control during the 1950s. It was felt that Integrated Pest Management should be primarily based on cultural and biological controls while chemical control should only be used as an 'emergency break' in those cases in which these natural controls failed. The concept of the economic damage threshold was coined to establish whether pest density was more damaging than the cost of treatment (Stern *et al.*, 1959). Methods were required to determine efficiently whether a pest population density was above or below this threshold. This decision problem requires a sampling methodology that results in a *classification* of density (Binns, 1994). A classification procedure may be designed such that the probability of a misclassification for densities deviating a specified amount from the threshold does not exceed a tolerance value (Binns, 1994; see below). Pests and diseases that pose a risk during a whole growing season may require repeated sampling through time.

Efficient monitoring can be achieved by linking classification procedures in time (Nyrop *et al.*, 1994; *cf.* second case study).

A special case of classification is *detection.* For instance, it may be asked whether a nematode species occurs in a piece of land or not. For a detection procedure, it is important to state explicitly the probability of a classification as nematode-free, when in fact the species is present (*cf.* third case study). This probability of misclassification decreases with nematode density. It depends also on the spatial aggregation of the nematodes and the sampling pattern. *Estimation* of density is often the appropriate objective in research (Wilson *et al.*, 1989), e.g. when describing the occurrence and dynamics of an organism in time and space. The *precision* of an estimate can be expressed as a standard error or as a variation coefficient. Density estimates (of sufficient accuracy) are inputs for *maps* showing spatial trends and patterns.

Table 1 purposes of sampling

technique	question/purpose
estimation	research on population dynamics or spatial pattern
detection	production of certified nematode-free seed potatoes
classification	intervene or not at a specific stage of pest phenology or crop growth
monitoring	intervene or resample later during a whole cropping season
mapping	visualizing spatial trends, patterns and relationships in pest density

One of the theoretical cornerstones of practical sampling programs in pest management is the description of the frequency distributions of observed pest and disease densities by means of statistical relationships (e.g. Taylors Power Law) and probability models (e.g. the negative binomial distribution). These mathematical tools are derived from an analysis of observations, in which the spatial dimension is neglected. This neglect is warranted when the interest of a grower or scout is in the overall density and its effect on crop productivity and quality and not in the spatial pattern. Some of the most important mathematical tools are presented. The use of these tools is illustrated in a case study on the development of a monitoring plan, based on sequential classification sampling plans.

Basic tools for describing sampling distributions

The spatial pattern, the sample size and the spatial distribution of samples affect the frequency distribution of sample counts (sampling distribution). If the location of each damaging organism were independent of that of the other, the spatial distribution would be random. For any size of sample and spatial arrangement of samples, the resulting frequency distribution can then be described by the Poisson probability

distribution, which is characterized by a single parameter, the average density (μ).

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

For the Poisson distribution, the variance of density is equal to the average density.

$$\sigma^2 = \mu$$

Subsequent probabilities of the Poisson distribution are calculated by

$$\begin{cases} P_0 = e^{-\mu} \\ P_{x+1} = \dfrac{\mu}{x+1} P_x \end{cases}$$

As a rule, however, pests and diseases deposit their offspring close to themselves, resulting in patchy distributions. Such spatial patterns result in frequency distributions with longer tails than the Poisson distribution. The negative binomial distribution is widely used (though not the only usable function) for describing these long-tailed frequency distributions. The negative binomial distribution is defined by:

$$P_x = \left(\frac{k}{k+\mu}\right)^k \binom{k+x-1}{x} \left(\frac{\mu}{k+\mu}\right)^x$$

The parameter *k* is called the dispersion parameter. The variance of the distribution and the length of the tail *decrease* with *k* (for given μ). For large *k*, the negative binomial distribution is similar to the Poisson distribution. For the negative binomial distribution, the variance of density is greater than the mean:

$$\sigma^2 = \mu\left(1 + \frac{\mu}{k}\right)$$

Probabilities of the negative binomial distribution are calculated with

$$\begin{cases} P_0 = \left(\dfrac{k}{k+\mu}\right)^k \\ P_{x+1} = \dfrac{k+x}{k+\mu} \dfrac{\mu}{x+1} P_x \end{cases}$$

The negative binomial distribution fits many observed frequency distributions because it has a quite flexible shape, ranging from the often bell-shaped low variance Poisson distribution (k -> ∞) to the monotonously decreasing high variance geometric distribution ($k = 1$) and beyond ($k < 1$).

In observed data sets, the parameter k usually has some relationship to the mean. This relationship can often be described with Taylors Power Law, which draws a linear relation between log(variance) and log(mean). For instance, for red mites on apple leaves, Nyrop & Binns (1991) used the relationship

$$\log\left(\hat{\sigma^2}\right) = \log(4.27) + 1.37 \log(\mu) \quad \text{or} \quad \hat{\sigma^2} = 4.27\,\mu^{1.37}$$

Using the relationship $k = \dfrac{\mu^2}{\sigma^2 - \mu}$

k can be estimated from the mean of the distribution.

Based upon a fitted probability distribution, the probability of the zero class can be used to estimate the relationship between the average density and the proportion of occupied sample units. This relationship can also be fitted with an empirical relationship

$$\ln\left(-\ln\left(1-p_T\right)\right) = a + b\ \ln(\mu)$$

Here p is the proportion of sample units with more than T specimens of the damaging organism, μ is average density, and a and b are regression parameters.

Sequential classification sampling

For the question whether a pest density or incidence is below or above a threshold, Walds Sequential Probability Ratio Test (SPRT) provides an optimal decision procedure. Instant recipes (and spreadsheets that do the calculations) are available to construct an SPRT-based sampling procedure for a range of sampling distributions, including the Poisson, negative binomial, binomial and normal distribution (Fowler & Lynch, 1987). For the negative binomial distribution, the following information is required and sufficient to construct an SPRT.

1. Information defining sampling performance
   The probability α of erroneously deciding that the average density is above the threshold T, when the actual density is in reality μ_0, which is smaller than T.
   The probability β of erroneously deciding that the average density is below the threshold T, when the actual density is in reality μ_1, which is greater than T.
2. Information defining the sampling distribution
   The dispersion parameter k

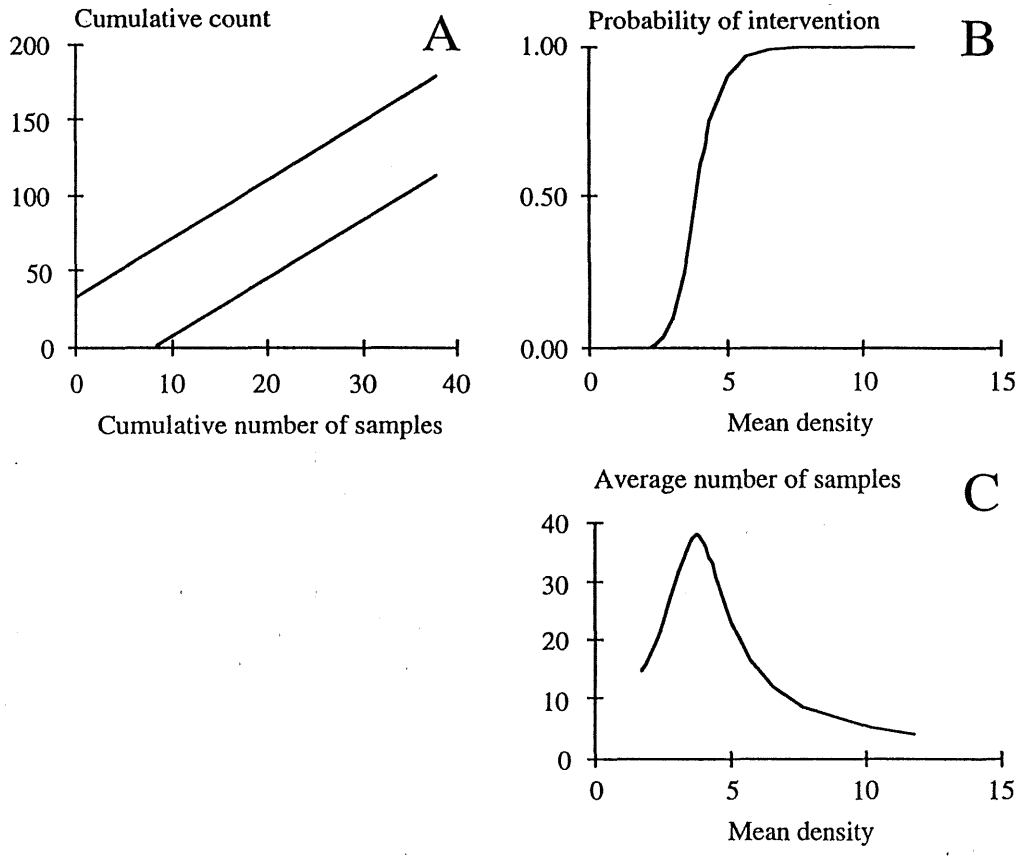

Fig. 2 (A) Stoplines of a sequential sampling plan, based on Walds Sequential Probability Ratio test; (B) Probability of intervention, as a function of the mean density; (C) Average number of samples required to make a decision, as a function of the mean density. This specific example was generated using a negative binomial distribution with $k = 0.6$; $\mu_0 = 3.0$; $\mu_1 = 5.0$; $\alpha = \beta = 0.1$.

The sequential plan is executed by inspecting sample units one by one and plotting the cumulative count against the cumulative number of sample units in Fig. 2A. When one of the two stoplines is crossed, sampling is terminated. When the higher stopline is crossed, the decision is to intervene. When the lower stopline is crossed, the decision is not to intervene.

The performance of a sequential sampling plan is judged by the average number of samples and the probability of intervention, which are both functions of the true mean density. The probability of intervention (Fig. 2B) is an increasing function of density, with the 50% point close to the threshold *T*. The average number of samples (Fig. 2C) is a maximum function. The highest number of samples is required (on average) when

the actual density is close to the threshold, because the probability of remaining between the stoplines is then greatest. The expected number of samples decreases as true density is further removed from the threshold. The two probability statements that define the SPRT, mark the position of two points on the probability of intervention-function of Fig. 2B: namely the points (μ_0, α) and $(\mu_1, 1-\beta)$. As a rule, μ_0 and μ_1 are chosen such that the threshold T is the average of them, while the error rates α and β are equal.

Steps in developing a monitoring protocol

When a pest must be monitored during an extended period of time, a procedure is required that ensures timely intervention when required, but that at the same time limits the frequency of observations as much as warranted. As an example of the practical application of probability and dynamical models, a guideline for developing a monitoring protocol is given. Similar guidelines may be developed for other sampling techniques, as classification, detection or estimation.

Step 1 Collect a data set defining the range of possible spatial distributions (irregularity in space, patchiness, variability), population dynamics (outbreaks, steady state, biological control) and the effect of the pest on growth and yield. This is the basic data set.

Step 2 Describe the sampling process, population dynamics and pest damage with mathematical models. These are basic models that serve as tools in the construction and evaluation of the monitoring protocol.

Step 3 Devise a monitoring protocol, using the basic models to take account of the spatial distribution, population dynamics and growth reducing effect of the pest or disease. The monitoring protocol provides decision support on when to make observations and whether or not to intervene.

Step 4 Simulate usage of the monitoring protocol. Calculate performance characteristics taking account of uncertainties in the outcome of sampling and in the dynamics of the pest by using stochastic parameters in the basic models.

Performance characteristics of a monitoring protocol are:

- total number of sampling occasions ± SD
- total number of samples ± SD
- overall probability of intervention
- cumulative pest density over time ± SD
- pest density at intervention ± SD

The first two variables are indices of effort while the others quantify the quality of control. Performance characteristics depend on spatial variability and pest dynamics

Step 5 Re-iterate steps 3 and 4, until an acceptable performance is attained. If no acceptable performance can be attained, consider developing alternative protocols for specific situations, for instance with and without natural enemies and for cold and warm weather.

Step 6 Test an acceptable monitoring protocol under field conditions.

Step 7 Apply the well-tested monitoring protocol in practice. Obtain feedback from users and continuously improve the protocol.

Variance components

A technique that is sometimes useful in defining sampling programs is analysis of variance and estimation of variance components. For instance, Nyrop & Binns (1991) discuss the contribution of between-tree-variation and within-tree-variation to total sampling uncertainty for a leaf miner species in apples. Based upon this analysis, it was concluded that it was more cost effective to inspect many trees and only few branches per tree than to sample many branches on a few trees. In the later case, the comparatively large variance component between-trees was not compensated for by adequate repetition. Based on an estimation of variance components and quantification of the costs of sampling trees and branches within trees, the most accurate sampling scheme for given cost could be defined. Alternatively, the cheapest scheme providing a minimum precision could be identified.

CASE STUDIES

Case: geostatistical analysis of black rot patterns in cabbage

The use of geostatistics is illustrated with data from black rot in cabbage. Black rot is caused by the bacterium *Xanthomonas campestris* pv. *campestris* (Pammel) Dowson 1939. The bacterium is seed-borne, and infested seed is an important source of inoculum. Cruciferous weeds, plant residue and cabbage volunteers are other inoculum sources. The pathogen can spread rapidly with wind and rain storms. The main objective of the study was to determine how far sampling intensity and observation time could be reduced, without obtaining a 'blurred' image of the spatial disease pattern.

A natural black rot epidemic was studied in a 20 by 20 m red cabbage field near Wageningen. The field had been planted with 1600 red cabbage plants in a 50 x 50 cm square arrangement on 15 May 1990. Rows ran northeast. Black rot incidence was scored visually on all 1600 plants on 24 August. This complete inventory is referred to as sampling plan I. To determine sampling with reduced intensity, geostatistical

mapping of the disease pattern in the plot was performed on the basis of reduced data sets. One set (referred to as sampling plan II) uses the presence/absence data on every third plant (yielding 533 data points). The third plan uses data on one in five plants (320 observations).

Semivariograms were calculated and fitted for each sampling intensity (Fig. 3). Exponential models gave the best fit. The calculated range for the three sampling intensities differed only slightly, varying from 2.7 with plan I to 2.2 m for plan III. An optimal sampling distance would be approximately 2.5 m.

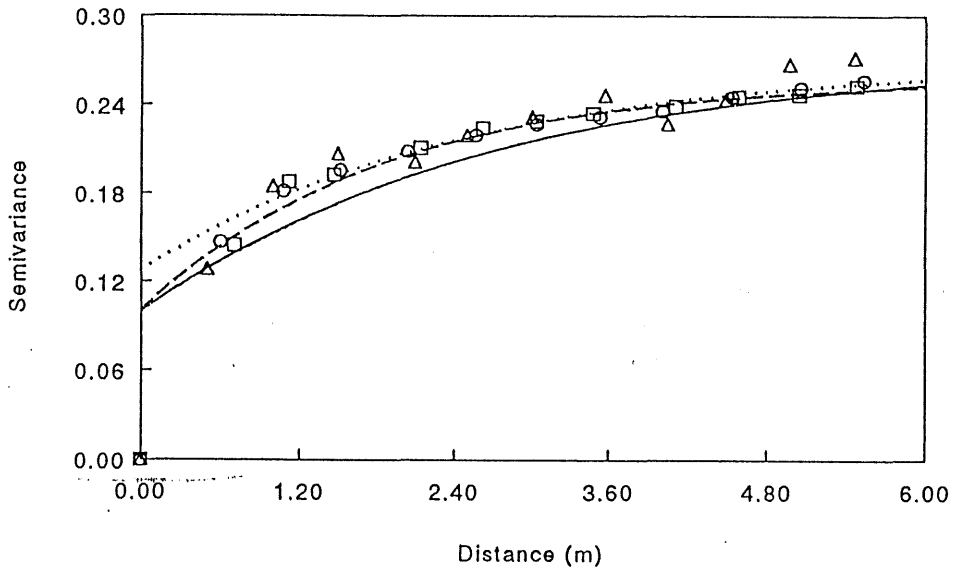

Fig. 3 Semivariograms of black rot disease incidence for sampling plan I (O ———; scoring all the 1600 plants), II (□ ·····; scoring one out of every three plants), and III (△ – – – – –; scoring one out of every five plants).

Directional semivariograms were calculated in northeast and southeast direction, to search for anisotropy resulting from predominant south-western winds during rainfall. Anisotropy was indeed found (Fig. 4). The semivariogram based on north-east distances had an eight times greater range (8.9 m) than the south east semivariogram (1.1 m). This result confirms the assumption that disease was spread with splashing rain storms blowing predominantly from the south west.

Disease incidence estimates for the three sampling plans were quite similar: 45.6, 43.1 and 44.9% for plans I-III, respectively. Kriging for plans II and III reproduced the observed disease incidence pattern accurately (Fig. 5). In this figure, each square represents the disease incidence per 4 plants. The distribution of black rot was not

homogenous; incidence was higher in the centre of the plot. Kriging based on plan II reproduced this pattern well; kriging based on plan III reproduced it less well, but the representation of the actual pattern is still acceptable.

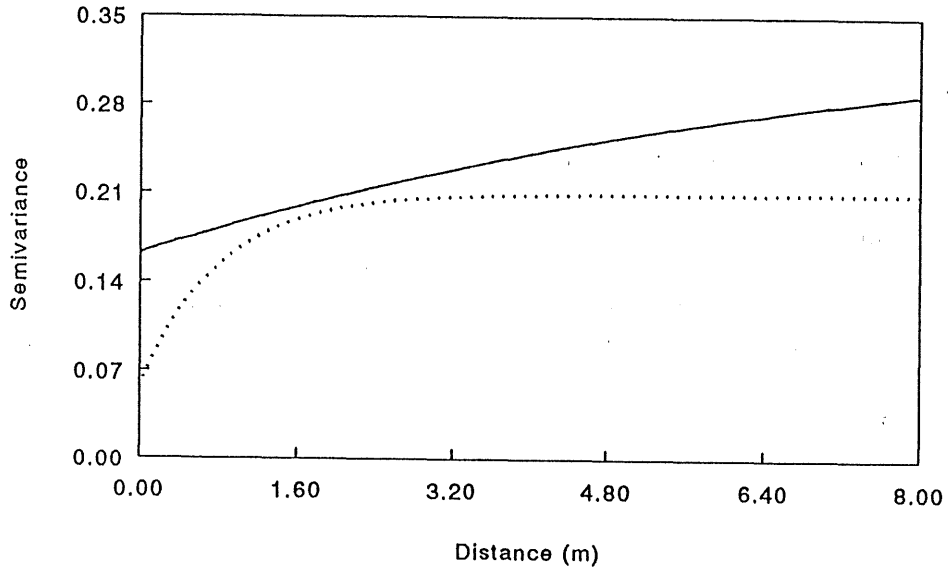

Fig. 4 Directional semivariograms of black rot disease incidence for sampling plan I. South west (prevailing wind): ———; North west (perpendicular to the wind): ----------

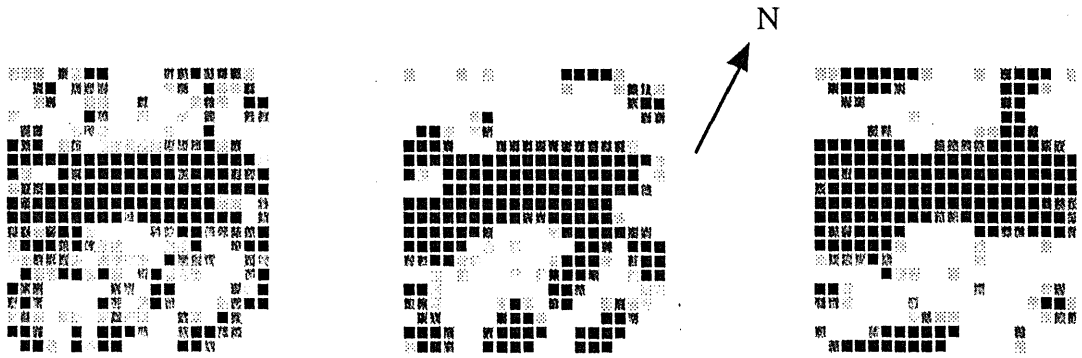

Fig. 5 Actual pattern of black rot incidence on 24 July 1990 and kriging maps based on sampling intensities of 33% (plan II) and 20% (plan III). Each square represents a quadrate of 2 x 2 plants. The intensity of shading indicates incidence in these quadrates, running from 0/4 infected plants (blank) to 4/4 infected plants (black).

This example illustrates the usefulness of geostatistics for analyzing and mapping spatial patterns. It was possible to reduce the number of samples with 80% and still obtain sufficient information about the spatial pattern. Sampling effort could probably not have been further reduced than this because the range of influence in the semivariogram was 2.5 m. When spatial correlation extends further, greater savings in sampling effort are attainable. For instance, Lecoustre *et al.* (1989) found that a sample size of 7% of all plants sufficed to assess the spatial pattern of African cassava mosaic virus.

Case: developing and evaluating protocols for mite monitoring

Fruit tree red spider mite, *Panonychus ulmi*, is a potential pest in apples worldwide. It can be controlled naturally by predators, but biological control is easily upset by pesticides. In apple crops in the state of New York, monitoring from early June to late August is required to make sure that biological control is effective. Schemes for efficient monitoring over time in this system were developed and evaluated by Nyrop & van der Werf (1994) and Nyrop *et al.*, (1994). The approach can be readily transferred to other systems.

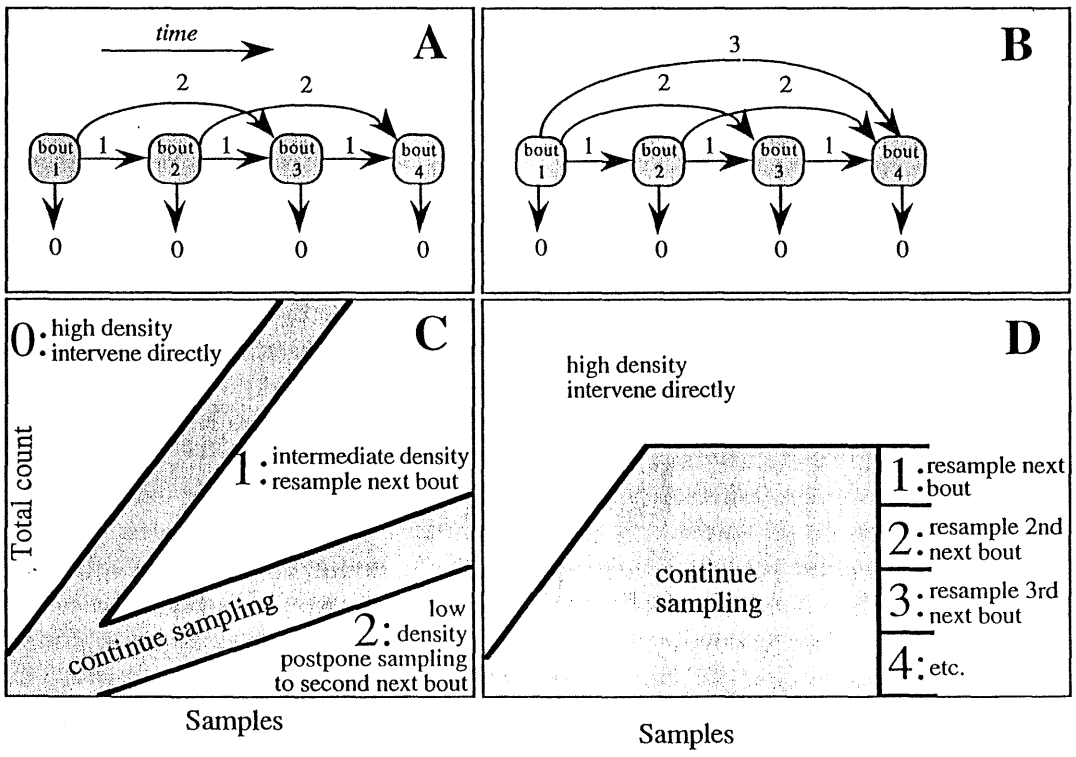

Fig. 6 Mite monitoring protocols (A & B) and the constituent sequential sampling plans (C & D). For explanation see text.

One method (cascaded tripartite sequential classification; TSC; Fig. 6A) was constructed by serially combining in time sampling plans that classify density into one of three categories with according management consequences 0 (intervene), 1 (sample at next occasion), and 2 (sample at second next occasion; Fig. 6C). The other protocol (adaptive frequency classification; AFC; Fig. 6B) was constructed by cascading in time sampling plans that are based on a combination of sequential classification and estimation of density (Fig. 6D). AFC allows sampling to be postponed more than two periods when density is unlikely to grow to damaging levels within that time.

In both schemes, timing and frequency of sampling are adjusted to the demands and possibilities of the actual situation, as indicated by sampling observations and a prediction of dynamics. The most suitable parameters for both schemes were found by simulating monitoring performance for fictitious and historical pest population trajectories. According to simulation, both methods scheduled interventions at appropriate times. The simulation results for the monitoring based on tripartite sequential classification were confirmed in a field evaluation involving 42 orchard blocks. Both methods use fewer sampling resources than sampling at pre-defined times, which is the usual method in practice. AFC-based plans required less sampling than TSC-based plans. Simulation further indicated that the currently used action threshold for red mites in the North Eastern USA are too low, resulting in spray recommendations when there would still be opportunity for natural control by predatory mites. TSC and AFC provide a framework for objectively evaluating and optimizing monitoring protocols for a range of pests and diseases.

Case: developing a detection method for nematode patches

Potato cyst nematodes (PCN) are not indigenous to Europe. They were introduced together with the potato plant from Central and South America. The presence of PCN manifested itself only in the 20th century when potatoes were grown in narrower rotations. Fields are free of PCN until an initial introduction occurs, mostly by seed potatoes. After introduction the nematode multiplies every year in which a host is grown. Active mobility of the nematode occurs after egg hatch, when the juveniles search for root tips to penetrate. This active movement would result in a dispersal of only a few centimeters per year. After maturing, the new generation of PCN overwinter as eggs inside the hardened dead body (cyst) of the female. Therefore, PCN are concentrated at locations where plant have grown. Horizontal and vertical redistribution of nematodes in the soil depends upon farming practices as soil tillage and harvesting. Dispersal from field to field occurs by pure chance when clumps of soil with cysts adhere to agricultural machinery or harvested potatoes.

In the newly reclaimed polder areas of the Netherlands, PCN infestations are young. They occur in the form of distinct patches in otherwise uninfested fields. Detection

methods for these PCN patches are being used in legislation, quarantine, certification of potatoes destined for export, and (most important) for guiding nematode control, e.g. by growing resistant cultivars. There was a need for estimating the error rates of the detection methods and to optimize (if possible) these methods.

As a first step, the shape of nematode patches was studied. About 40 farmer's fields, which, according to the statutory soil sampling protocol, were regarded as PCN infested, were sampled twice. The first sample was used to locate the infestation focus. The second sample was aimed at accurately mapping the spatial distribution of the cysts in the focus. Soil samples of at least 1.5 kg per m^2 were collected and processed.

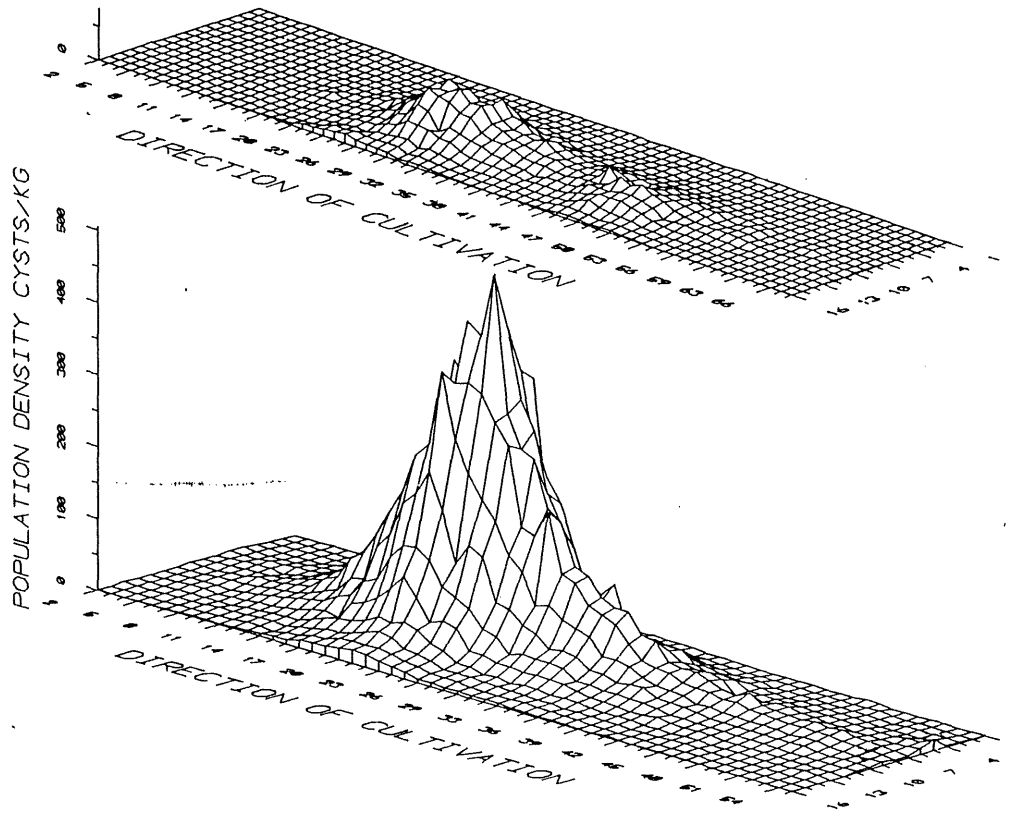

Fig. 7 Two foci of potato cyst nematode on heavy marine clay soil in a recently infested area. Above: small focus with a population density of 85 cysts per kilogram soil in the center of the focus. Below: large focus with more than 500 cysts per kilogram soil in the center of the focus. Both foci were mapped by sampling each square meter and collecting 2.5 kg of soil per m^2.

All foci were more or less elliptical with the largest population densities in the center. From this point the densities decreased exponentially. The decrease was slower in directions parallel to the rows than across them (Fig. 7), i.e. the patches had the greatest extension along the path of the machinery. The spatial extension of foci was also greater in the driving direction than in the reverse direction. (Because machinery has standard width, farmers may year after year follow the same driving pattern and directions over the field.)

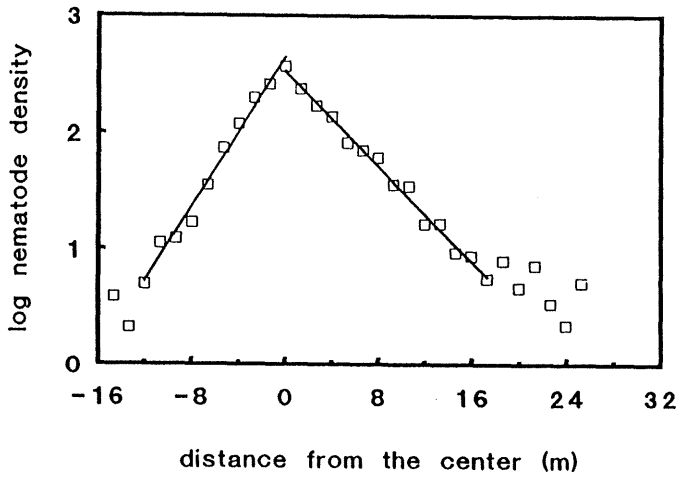

Fig. 8 Linear relationship between logarithm of nematode density and distance from the center of the focus. Squares represent actual cyst counts in samples of 2.5 kg soil, originating from the middle row of successive square meter plots in the direction of cultivation of the large focus depicted in Fig. 7. Samples with fewer than 5 cysts per kilogram soil were omitted when fitting the drawn regression line.

A linear relation between log population density and the distance from the center of a focus (Fig. 8) was found and parametrized (Schomaker & Been, 1992):

$$\mathrm{E}\left(N_{x,y}\right) = N_{0,0}\, L^{x}\, B^{y}$$

where

$\mathrm{E}\left(N_{x,y}\right)$	is the expected density at location (x,y)
(x,y)	is the location relative to the centre of the focus, with x measured in the direction of cultivation and y measured across
$N_{0,0}$	is the density in the centre of the focus
L	is the fractional decrease of expected density per meter departure from the centre along the rows, i.e. in the x direction

B is the fractional decrease of expected density per meter departure from the centre across the rows, i.e. in the y direction.

The frequency distributions of L and B in the 40 fields were approximately normal.

The frequency distribution of numbers of cysts in 1.5 kg samples from small areas as used for mapping the foci, is adequately described by a negative binomial distribution with a value of 70 for the parameter k. The probability of finding no cysts in a sample is then given by the zero-term of the negative binomial distribution.

An infestation focus is detected if one or more cysts are extracted from it. The detection probability of a focus can therefore be defined as 1 minus the probability that cysts were found in none of the subsamples taken from the focus. In the Netherlands and in most other countries sampling according to a rectangular grid pattern is customary. The distance from one core to the next in both directions determines how many subsamples are taken from a certain area. Sampling grid and auger size determine the total amount of soil collected. A sampling grid can be superimposed on a focus in many ways. Each overlay pattern of focus and sampling grid results in a different detection probability. A computer program was written to calculate the detection probability when shifting the sampling grid longitudinally and laterally, relative to the focus.

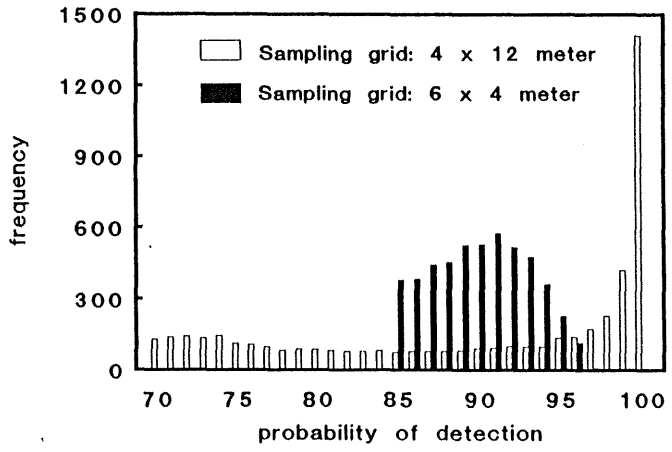

Fig. 9 A comparison of the frequency distributions of detection probability using a 6 x 4 m and a 4 x 12 m sampling grid (length x breadth). The core size was optimized to obtain an average detection probability of 90% for a nematode focus with 50 cysts per kilogram soil in the center.

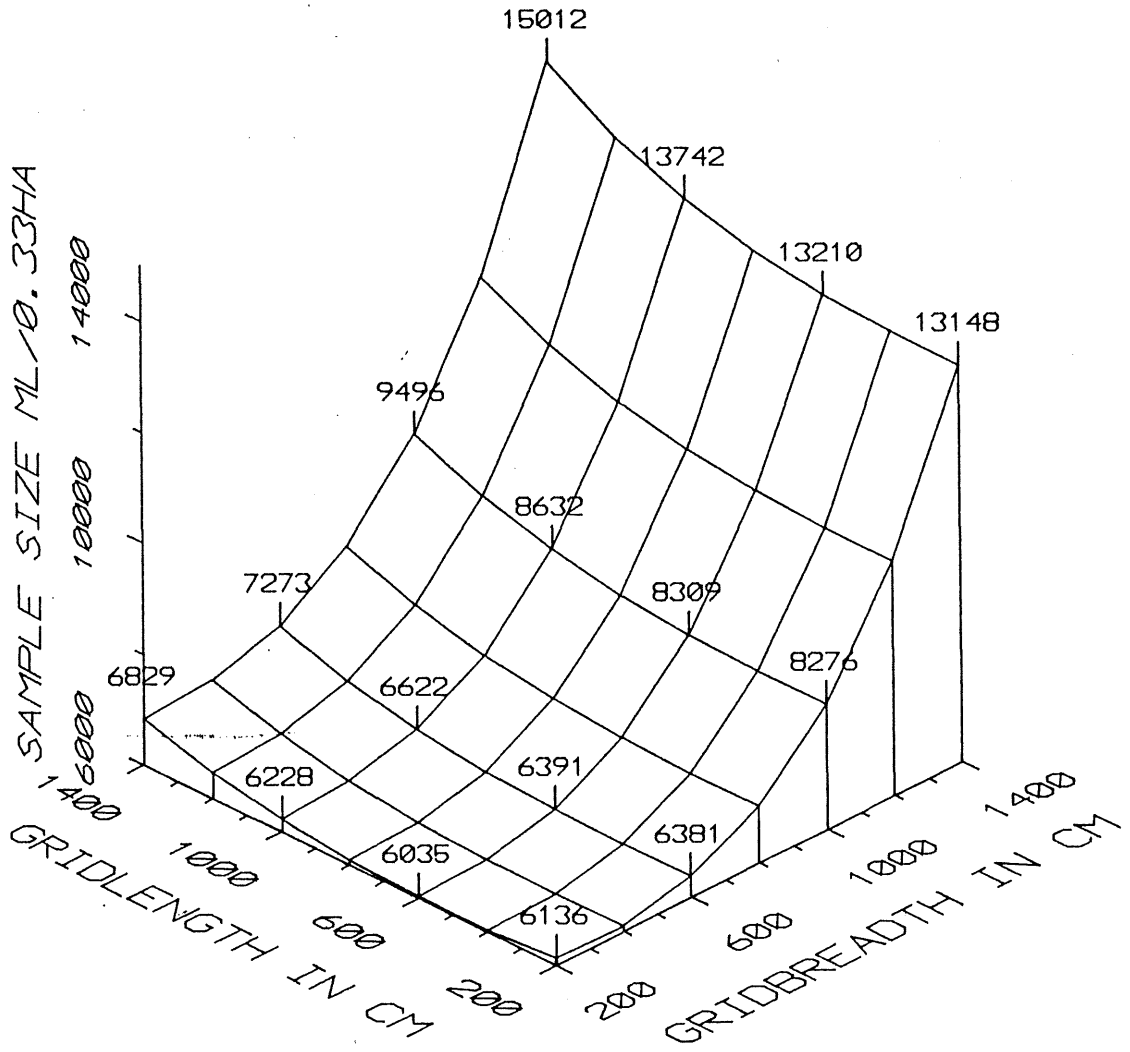

Fig. 10 A comparison of the sample size per 0.33 ha when the sampling grid was varied between 2 x 2 m and 14 x 14 m. For each grid, the core size were determined that yielded a 90% chance of detecting a single focus with 50 cysts per kg soil in the center. The total sample size is the product of core size and the total number of samples as determined by the grid.

Fig. 9 shows the frequency distributions of detection probabilities of two different sampling grids. The program uses the exponential equation that describes the spatial profile of foci to calculate the expected population densities throughout a predefined focus. The probability of detecting no cysts at a certain location in the focus was calculated with the negative binomial distribution. As detection with a low failure rate is required, the parameters L and B were set to relatively small values (10% percentiles of the observed distributions of L and B), reflecting steep gradients. This

'worst case approach' ensures that for 90% of the patches, the specified detection probability is actually attained. For 10% of the patches (the steepest ones), the detection chance will be lower than specified, and the overall detection chance for the whole population of patches will be better than specified. Calculations for a given sampling grid were made of the amount of soil per core and in total that had to be collected for a 90% detection probability of the predefined focus.

In Fig. 10, the required total sample sizes (g soil) per one third of a hectare are compared, when using different sampling grids. The narrowest grid had sampling intervals of 2 x 2 m; the widest grid 14 x 14 m. Iterations were made for sample size until, with each grid, a focus of 50 cysts per kg soil in the center was detected with 90% probability. The optimal sampling grid (in terms of the required amount of soil) was the 6 x 4 m grid (length x breadth), which required a total sample of 6 kg soil. To obtain the same 90% detection chance with a 4 x 12 m grid, 11 kg soil had to be collected and analysed.

Potato cropping frequencies differ among growing areas in the Netherlands, and among European countries. It is likely that the differences in cropping practices are reflected in spatial patterns of PCN, which affects the performance of sampling patterns. The required detection probability depends upon the product. Seed potato growers, always alert with regard to export requirements, require more precise detection methods than consumption potato growers. As a result of differences in spatial patterns and required detection probability, tailor-made sampling methods are desirable. The presented approach allows the design of sampling methods that are tailor-made for the respective target areas and product groups.

EPILOGUE

This paper draws together some techniques that are used in the study of plant pest and diseases, and could be profitable in the young research field of entomopathogenic nematodes. Progress in ecological research on EPNs is presently hampered by technical difficulties in retrieving nematodes from soil and by lack of knowledge on spatial patterns and sampling distributions. For entomopathogenic nematodes, most of the basic work on spatial patterns and sampling distributions that is necessary for the design of efficient sampling plans has still to be done. It is hoped that this presentation of research on pests and diseases provides ideas and stimulus to undertake such work and provide a sound basis for further ecological work.

The selection of techniques in this paper is necessarily restricted. A comprehensive treatise of techniques for spatial statistical analysis is given by Upton and Fingleton (1985; 1989). No mention is made here of techniques for modelling spatial *processes*. Such spatial modelling may explain spatial phenomena in relation to the

causal relationships and processes, contrary to statistical analyses, that can only describe and quantify correlations and trends, but cannot explain them. Overviews of approaches to spatial modelling are given by van der Werf *et al.* (1989) and Holmes *et al.* (1994). A potentially relevant new development are techniques of precision farming or site-specific management. These techniques are targeted at providing the appropriate management action to each site in a field, e.g. in response to soil fertility level (Wollenhaupt *et al.*, 1994; Bouma *et al.*, 1995) or weed development (Christensen *et al.*, 1994).

ACKNOWLEDGEMENT

We are indebted to Jan Nyrop, Michael Binns, Corrie Schomaker and Herman Frinking, with whom we collaborated on the presented case studies. They contributed substantially to the presented results and ideas.

REFERENCES

Binns, M.R., 1994. Sequential sampling for classifying pest status. In: L.P. Pedigo & G.D. Buntin: Handbook of sampling methods for arthropods in agriculture. CRC Press, Boca Raton, p. 138-173.

Bouma, J., J. Brouwer, A. Verhagen & H.W.G. Booltink, 1995. Site specific land management on a field level: high and low tech approaches. In: Conference Proceedings on Eco-regional Approaches for Sustainable Land Use and Food Production, Den Haag, Netherlands, 12-16 December 1994.

Burrough, P.A. 1987. Spatial aspects of ecological data. In: R.H.G. Jongman, C.J.F. ter Braak & O.F.R. van Tongeren (Eds): Data analysis in community and landscape ecology. Pudoc, Wageningen, p. 213-251.

Chellemi, D.O., K.G. Rohrbach, R.S. Yost & R.M. Sonoda, 1988. Analysis of the spatial pattern of plant pathogens and diseased plants using geostatistics. Phytopathology 78, 221-226.

Christensen, S., E. Nordbo & K. Kristensen, 1994. Weed cover mapping with spectral reflectance measurements. In: P. Brain, S.H. Hockland, P.D. Lancashire & L.C. Sim (Eds): Sampling to make decisions. Aspects of Applied Biology 37, p. 171-179.

Fowler, G.W. & A.M. Lynch, 1987. Sampling plans in insect pest management based on Wald's sequential probability ratio test. Environmental Entomology 16, 345-354.

Holmes, E.E., M.A. Lewis, J.E. Banks & R.R. Veit, 1994. Partial differential equations in ecology: spatial interactions and population dynamics. Ecology 75, 17-29.

Hudelson, B.D., M.K. Clayton, K.P. Smith, D.I. Rouse & C.D. Upper, 1989. Nonrandom patterns of bacterial brown spot in snap bean row segments. Phytopathology 79, 674-681.

Journel, A.J. & Huijbregts, C.J., 1978. Mining geostatistics. Academic Press, New York, 600 pp.

Lannou, C. & S. Savary, 1991. The spatial structure of spontaneous epidemics of different diseases in a groundnut plot. Netherlands Journal of Plant Pathology 97, 355-368.

Lecoustre, R., D. Fargette, C. Fauguet & P. Reffye, 1989. Analysis and mapping of the spatial spread of African cassava mosaic virus using geostatistics and the kriging technique. Phytopathology 79, 913-920.

Munkvold, G.P., J.A. Duthie & J.J. Marois, 1993. Spatial patterns of grapevines with Eutypa dieback in vineyards with or without perithecia. Phytopathology 83, 1440-1448.

Nelson, M.R., R. Felis-Gastelum, T.V. Orum, L.J. Stowel & D.E. Meyers, 1994. Geographical information system and geostatistics in the design and validation of regional plant virus management progress. Phytopathology 84, 898-905.

Nyrop, J.P. & M.R. Binns, 1991. Quantitative methods for designing and analyzing sampling programs for use in pest management.In: D. Pimentel (Ed.): CRC handbook of pest management in agriculture, p. 67-132.

Nyrop, J.P., M.R. Binns, W. van der Werf & J. Kovach, 1994. Cascading tripartite binomial classification plans to monitor European red mite (Acari, Tetranychidae) through a season; development and evaluation of a new methodology for pest monitoring. Experimental & Applied Acarology 18, 123-153.

Nyrop, J.P. & Van der Werf, W., 1994. Sampling to predict or monitor biological control. In L.P. Pedigo and G.D. Buntin (Eds.): Handbook of Sampling Methods for Arthropods Pests in Agriculture. CRC Press, Boca Raton, Florida. p. 245-336.

Stein, C.G. A., Kocks, J.C. Zadoks, H.D. Frinking, M.A. Ruissen & D.E. Myers, 1994. A geostatistical analysis of the spatio-temporal development of downy mildew epidemics in cabbage. Phytopathology 84, 1227-1239.

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

Schomaker, C.H. & T.H. Been,1992. Sampling strategies for the detection of potato cyst nematodes: developing and evaluating a model. In: F.J. Gommmers & P.W.Th. Maas, (Eds): Nematology from molecule to ecosystem. Proceedings of the Second International Nematology Congress, 11-17 August 1990, Veldhoven, the Netherlands. Dekker & Huisman, Wildervank, the Netherlands. p. 182-194.

Upton, G.J.G. & B. Fingleton, 1989. Spatial data analysis by example. Volume 1. Point pattern and quantitative data. Wiley, Chichester, 410 pp.

Upton, G.J.G. & B. Fingleton, 1989. Spatial data analysis by example. Volume 2. Categorical and directional data. Wiley, Chichester, 416 pp.

van der Werf, W., W.A.H. Rossing, R. Rabbinge, M.D. de Jong & P.J.M. Mols, 1989. Approaches to modelling the spatial dynamics of pests and diseases. In: R. Cavalloro and V. Deluchi (Eds): PARASITIS 88, Proceedings of a scientific

congress, Barcelona, 25-28 October 1988, p. 89-119.
Wilson, L.T., W.L. Sterling, D.R. Rummel & J.E. DeVay, 1989. Quantitative sampling principles in cotton IPM. In: Frisbie, R.E., K.M. El-Zik & L.T. Wilson (Eds): Integrated pest management systems and cotton production. John Wiley & Sons, New York, p. 85-119.
Wollenhaupt, N.C., R.P. Wolkowski & M.K. Clayton, 1994. Nutrient management; mapping soil test phosphorus and potassium for variable-rate fertilizer application. Journal of Production Agriculture 7, 441-448.