Analysis and Mapping of the Spatial Spread of African Cassava Mosaic Virus Using Geostatistics and the Kriging Technique

R. Lecoustre, D. Fargette, C. Fauquet, and P. de Reffye

First and fourth authors, Laboratoire de Modélisation, GERDAT/CIRAD, Centre de Recherches de Montpellier, BP 5035 34032, Montpellier cedex, France; second and third authors, Laboratoire de Virologie, ORSTOM Adiopodoumé, BP V 51, Abidjan, Ivory Coast.

We thank J. M. Thresh and P. Nicot for helpful discussions and constructive criticism of the manuscript.
Accepted for publication 10 October 1988.

ABSTRACT

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

Theories of regionalized variables and kriging were used to assess the spatial pattern of African cassava mosaic virus (ACMV). A linearlike semivariogram without a range characterizes the ACMV distribution and indicates a strongly spatially dependent structure with limited random variation. Oriented semivariograms reveal a strong anisotropy in relation to the prevailing wind direction. Further features of the semivariogram and comparisons of semivariograms between fields and between surveys provide additional information and support various hypotheses on the pattern of spread. From a sample of limited size, kriging reproduced the main characteristics of the spatial pattern of spread, including higher incidence along the wind-exposed southwest field borders, disease gradients, and other less obvious features. Up to 60% of the total variance was reconstructed from a 7% sample. Kriging was successfully applied to characterize the spatial pattern of spread in cassava fields differing in planting date, size, arrangement, orientation, and method of sampling. This technique was also efficient when the pattern of spread was heterogenous, although more intensive surveys were then required. Practical applications of geostatistics and kriging in epidemiology are discussed.

Spatial patterns of disease level can provide important clues to the ecology of disease (e.g., direction and distance of spread, importance and proximity of the sources of virus and vectors, vector mobility) (20). Patterns must be considered in designing sampling methods and sound control measures (8,20). Hence, efficient methods of analysis and interpretation of spatial pattern are needed to provide the greatest possible information in relation to the time and effort involved. Past studies of spatial patterns have relied mostly on methods based on the examination of the mean and variance or on the frequency distribution of observed disease incidence (17). However, these methods do not incorporate information on the location of the samples and in particular they fail to consider the degree of dependency between neighboring observations (i.e., spatial dependence). Recently, methods that recognize such dependency have been proposed (4,18,19). Geostatistics, as introduced by geologists, quantifies the spatial dependence and has been applied successfully in agroforestry, agronomy, and entomology (3,12,13,16). It has been proposed recently to analyze the spatial spread of plant pathogens (4,12). Geostatistics uses the theory of regionalized variables and only requires an assumption that the variance between samples is a function of the distance of separation. ("Semivariance," as defined subsequently, is a measure of the expected squared difference between all values separated by the same distance.) The semivariogram plots the semivariances versus distance and illustrates the spatial variation.

Monitoring the incidence and spread of plant virus diseases requires extensive and repeated surveys, which are time-consuming, expensive, and inconvenient. Furthermore, very large plantings may be impracticable to survey. An efficient way of mapping the spatial pattern of spread, based on a sample of limited size but able to reconstruct the main characteristics of disease distribution, would be a useful tool in phytopathology. Kriging is such a technique; it makes optimal, unbiased estimates of regionalized variables at unsampled locations using the structural properties of the semivariogram and the initial set of data values (9). The technique has been widely used for mapping in geology and pedology (21).

African cassava mosaic disease is caused by a whitefly-borne geminivirus (2). The spatial patterns of spread of this disease have been studied intensively in the Ivory Coast (5,6) and are mainly characterized by gradients oriented in the direction of the prevailing southwest wind. In this article, the theory of regionalized variables is used to assess the spatial patterns of the spread of African cassava mosaic virus (ACMV) in various cassava fields that differ in total area, subplot size, planting dates, and orientation. We also describe the application of kriging to reconstruct the spatial patterns of spread of ACMV within plantings, using data from a limited number of sample points.

MATERIALS AND METHODS

The theory of regionalized variables. A variable is "regionalized" when its values depend on its spatial position (14). A simple example (15) illustrates this concept. Two series of measurements made of a hypothetical variable at regular intervals along a row in a field gave the following numerical sequences:

A: 1-2-3-4-5-6-5-4-3-2-1,
B: 1-4-3-6-1-5-3-4-2-5-2.

Sequence A has a clearly defined symmetry, whereas any structure for sequence B is irregular and difficult to define. Nevertheless, the two series of 11 measurements have the same mean and variance; thus it is impossible to adequately describe the detailed spatial distribution of the variable by using only these two parameters. A regionalized variable arises from the combination of two contrasting aspects. The first is a random effect, as the studied variable presents irregularities in space that are not predictable from point to point. The second is a structural aspect, characteristic of a regionalized phenomenon, where the data are organized in space. Mineral content, water resource, insect numbers, and disease incidence may be considered regionalized variables.

Semivariograms. Geostatistics detects spatial dependence by measuring the variation of regionalized variables among samples separated by the same distance. The semivariance is the average

of the squared differences in values between pairs of samples separated by a given distance h. The analytical tool is a semivariogram $G(h)$, which plots the semivariance versus distance. It is defined for any distance h:

$$G(h) = [1/(2N_h)] \Sigma[F(x_i + h) - F(x_i)]^2 ,$$

where x_i is the position of one sample of the pair, $x_i + h$ is the position of another sample h units away, $F(x)$ is the measure of a value at location x_i, and N_h is the number of pairs (x_i, $x_i + h$). When the distance becomes great, the sample values may become independent of one another and then $G(h)$ tends towards a maximum value. The value a of h, corresponding to this maximum, is called the semivariogram range and corresponds to the distance at which correlation between values taken at the sample points is negligible.

The shapes of the experimental semivariograms may be highly variable. The semivariogram immediately takes its maximum value if there is no correlation and signifies that the phenomenon is completely random. It is represented by a flat semivariogram: the "pure nugget" effect. This depends on microstructure and is usually superimposed on other structures. The observed semivariogram can be adjusted to several theoretical models, including spherical, exponential, Gaussian, and linear (14). The linear model does not have a plateau and may be considered to be the beginning of the spherical or exponential model (14). Its equation is $G(h) = Go + bh$.

Anisotropy characterizes a regionalized variable that does not have the same properties in all directions. Semivariograms can be calculated for all directions combined or for specific directions to test for anisotropy. If the structure cannot be demonstrated in one particular direction, it suggests that the structure is oriented along an axis perpendicular to that direction (14).

Kriging. If the adjusted semivariogram describing a given variable for a selected model is known, a local estimate can be made of the regionalized variable from a sample collected experimentally. Kriging is the estimation method. This method is termed unbiased as, unlike other more simple methods, it plots the mean and variance of the phenomenon, restores the values measured at sample points, and ensures that the estimation variance is minimized. The size of the "window" defines the square area centered on the point to be estimated, the width of which maximally equals $\sqrt{2}$ times the practical range of the semivariogram. An estimate of the value $F(x_0)$ at any point x_0 surrounded by n points sampled, is a linear combination of experimental values.

$$F(x_0) = \Sigma L_i F(x_i) ,$$

where $F(x_i)$ designates, as before, the value of the variable at point x_i, and L_i is the weighted coefficient of the sample x_i. The L_i values are calculated with the modeled semivariogram (3) so that the expected variance value at point x_0 is minimum and with $\Sigma L_i = 1$. It is inappropriate to use sample points from distances greater than the semivariogram range to estimate the value $F(x_0)$ at any point. In the case of the linear semivariogram, there is no practical range. Then, the size of the window is limited only by the size of the smallest dimension of the field. At least two points are required within a window, as a single point leads to a linear estimation.

Field surveys. Analyses were made of data from three field trials at the Adiopodoume Experimental Station of ORSTOM, 20 km west of Abidjan, Ivory Coast. The plantings were of healthy cassava cuttings (cultivar CB) obtained from the Toumodi Experimental Station in the savannah region, 200 km north of Abidjan. Disease incidence was assessed in plots of 100 plants (arranged 10 × 10 at 1 × 1 m spacing) in fields 1 and 3 and in subplots of 25 (5 × 5 at 1 × 1 m spacing) in field 2. In each trial, disease incidence was assessed visually. Field 1 of 1.0 ha was planted in October 1982. Disease incidence was recorded every 2 wk for 8 mo. Diseased plants were labeled and left in place. The pattern of spread in field 1 is described in detail elsewhere and is typical of ACMV spread in large cassava plantings subject to edge effects (6). Field 2 was square, of 0.49 ha, planted in July 1983, and oriented with the upwind margin across the direction of the prevailing southwest wind. Disease incidence was recorded weekly in each plot, and diseased cassava plants were removed after they had been recorded. Disease incidence was assessed initially in the 196 subplots of 25 plants, then recalculated in the 49 plots of 100 plants by combining four adjacent subplots. Field 3, of 4.0 ha, was planted in October 1984 as four blocks of 1.0 ha, each separated by a path 3 m wide. Disease incidence was recorded in January 1985 in plots of 100 plants, and diseased plants were left in place.

Methodology. The first step was to analyze the experimental semivariogram and to fit a model. The validity of the fit was evaluated by calculating the correlation coefficient between observed semivariogram values and the model predictions. The nonoriented semivariogram was studied first. To analyze the anisotropy of the variable, we also studied the semivariograms oriented in four principal directions. The precision of the estimates depends not only on the quality of the adjustment between the observed semivariogram and the modeled semivariogram, but also on the density and distribution of the samples. Then, the second step was to determine the sampling characteristics: density and distribution of the samples and size of the window. The third step was to investigate whether the kriging technique used with the established sampling procedures could reproduce the observed pattern of spread within the different cassava fields. The calculated patterns of spread were then compared with those observed in fields 1, 2, and 3 by comparing the maps of spread and by calculating the correlation coefficient between calculated and observed values.

RESULTS

Experimental and adjusted semivariograms. Figure 1A presents the experimental semivariogram for field 1, 7 mo after planting. The experimental semivariogram could be fitted closely to a linear model ($r = 0.97$, df = 11). Such a linear relationship appears to be typical of ACMV spread in our experiments, as it was also observed in field 3 ($r = 0.93$, df = 25, Fig. 1B). Semivariograms for field 2 were calculated 6, 7, and 8 mo after planting, corresponding to increasing levels of infection. Disease incidence was calculated in plots of 100 plants. Each semivariogram could be described adequately by a linear model (Fig. 1C); correlation coefficients were 0.98, 0.99, and 0.98 (df = 6) at 6, 7, and 8 mo after planting, respectively.

The semivariograms exhibited several characteristics. All had nonzero semivariances as h tended towards zero. This is the "nugget variance" and represents unexplained or "random" variance. In the fields surveyed, nugget variance was limited, which indicates that the spatial pattern of ACMV spread had a strongly spatially dependent structure with limited random variation. Actually, with the linear model, it is the high ratio—slope of the regression line divided by nugget variance—that quantifies precisely the spatial component of the structure of the spread (R. Lecoustre, *unpublished*). In all fields, the semivariance increased continuously without showing a definite range. This indicates that the greater the separation of the samples, the greater the difference in disease incidence. However, the systematic deviations of the experimental points from the regression line indicated that the semivariance was not strictly proportional to the distance between points. These deviations further characterized the ACMV pattern of spread. For example, concavities observed for distances between points 40–60 m apart for field 1 (Fig. 1A) and 20–30 m for field 2 (Fig. 1C) were likely to be related to border effects, which were very pronounced at these distances in their respective fields. In field 3, a change of slope was observed for distances between points around 100 m (Fig. 1B), which may reflect the fact that this field consisted of four distinct blocks of 100 × 100 m each separated by a 3-m wide path with high incidence on each side of this path.

Oriented semivariograms. With oriented variables, the

semivariogram could be calculated in each direction to find a direction with higher degrees of autocorrelation. Figure 2A illustrates the semivariogram in the southwest-northeast direction in field 1, which fitted to a linear model. Results were similar for the north-south and east-west semivariograms (not illustrated). By contrast, Figure 2B illustrates the semivariogram along the northwest-southeast axis, which showed no pattern. The semivariogram along the northwest-southeast axis passed through the origin as h tended to zero. These semivariograms indicated a strong anisotropy of the variable and revealed a disease gradient effect oriented along the southwest-northeast diagonal, which was the prevailing wind direction. For practical reasons, due to the small number of sample points used to determine the experimental semivariogram, the nonoriented semivariogram had to be chosen. This was valid as the close agreement with the linear model indicates little perturbation due to the prevailing wind direction.

Sampling procedures. Various sampling procedures were used to survey plant virus diseases (1). In preliminary studies, using field 1, we tested sample sizes of 4, 7, 13, and 25% of the total stand with different window sizes at several dates corresponding to various amounts of spread. Table 1 presents results obtained from data collected 3 mo after planting. A 7% sample with a window of size 9 gave a good correlation between observed and calculated patterns of spread (Table 1). Several random 7% samples were tested, as well as one taking into account the four "corner plots" (Fig. 3). Correlation coefficients between observed and calculated mapping of the various random samples ranged from 0.51 to 0.82 (df = 98). This variation indicated that the position of the samples was critical for efficient mapping. A value of 0.81 was obtained for the sample that included the corner plots; this value did not differ significantly from the 0.82 value drawn from the most efficient random pattern of sampling. Similar tests at other dates and in other fields confirmed that samples that include the corner plots provided the best correlations between the observed and calculated values. Subsequently, a sampling pattern that specifically included the corner plots was applied in the following analyses.

Mapping. In field 1, a close correlation was found between the observed and calculated values ($r = 0.78$, df = 98) as illustrated in Figure 3. In addition, kriging allowed a good reconstruction of the observed mapping using a sample of limited size; Figure 4 illustrates the observed and calculated distributions of disease 7 mo after planting as based on a 7% sample. Infection was not homogenous throughout the field, and the wind-exposed south and west borders had a higher disease incidence than the north and east borders and also than the center of the field. Kriging gave a calculated pattern of disease closely related to the one

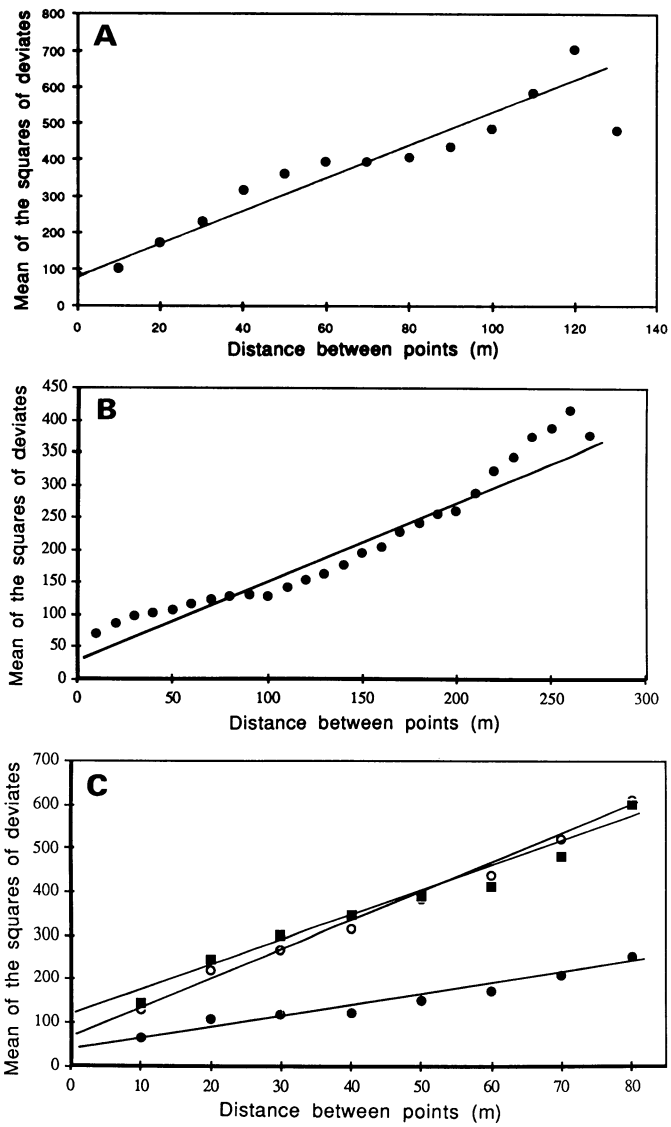

Fig. 1. Nonoriented semivariograms of cassava disease spread. **A,** Field 1, 7 mo after planting. **B,** Field 3, 4 mo after planting. **C,** Field 2 at 6 (●), 7 (○), and 8 (□) mo after planting.

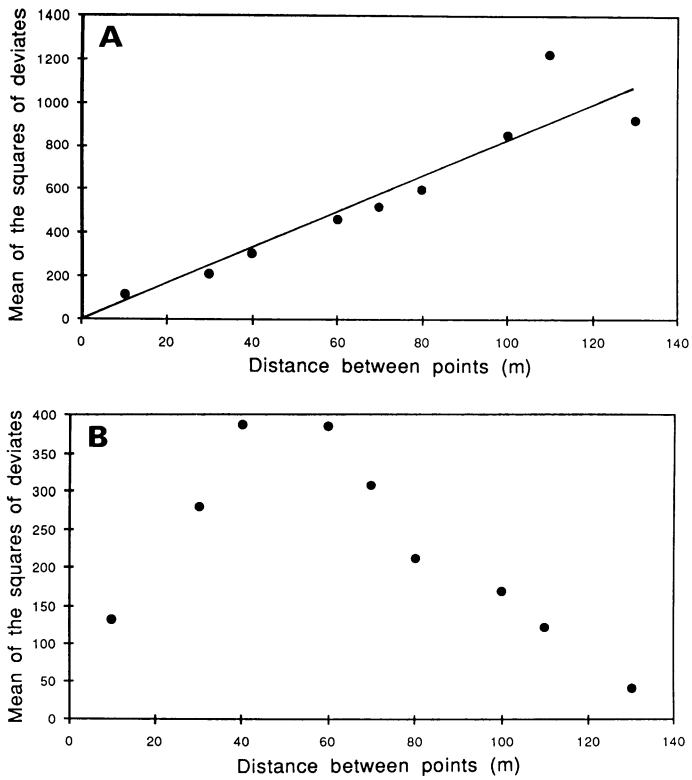

Fig. 2. Oriented semivariograms of cassava disease along the southwest-northeast axis (**A**) and along the southeast-northwest direction (**B**) 7 mo after planting in field 1.

TABLE 1. Correlation coefficients between calculated and observed mappings for different sample sizes and different window sizes for field 1, 3 mo after planting

Percent of sampling	Window size			
	3	5	7	9
4	*[a]	*	0.49	0.56
7	*	*	0.79	0.82
13	*	0.76	0.76	0.60
25	0.86	0.86	0.07	−0.19

[a] Asterisk indicates an impossible combination, as a minimum of two points is required.

observed (r = 0.78, df = 98) and reproduced higher incidence at the upwind field borders. Disease incidence in the 36 blocks of the first two exposed borders ranged from 60 to 100%, and all but one of the calculated values for these blocks fell within this range. The observed disease gradient along the southwest-northeast axis was characterized by a sharp decrease of disease incidence from nearly 100% along the upwind edges to 30% at the center of the field, followed by an increase towards 50% at the downwind borders (Fig. 5A). The disease gradient reproduced by kriging (Fig. 5A) closely paralleled the one observed (r = 0.97, df = 8). Along the southeast-northwest axis also, the general pattern of disease incidence was reproduced (r = 0.82, df = 8) (Fig. 5B).

The distribution of disease within field 1 was typical of that usually found for ACMV (6). It was of interest to determine how efficiently kriging reproduces the pattern of spread in fields differing from field 1 in size, degree of exposure to wind, disposition, and overall disease incidence. In fact, the spatial pattern of spread within field 2, although showing the greatest incidence of disease along the wind-exposed border, differed somewhat from that observed in field 1. The disease gradient was less clearly marked, and the overall "background" incidence of disease was reached only 20–30 m from the border. However, there was good agreement between the observed incidence of ACMV and the incidence calculated from a 7% sample comprising 14 of the 196 separate subplots of 25 plants (r = 0.76, df = 194) (Fig. 6). Disease incidence was higher in the first five rows of plots along the southwest border than elsewhere and decreased with increasing distance from this border. The lowest incidence was in the middle of the field. Additional features of the observed spread such as a slight increase in disease incidence on the northeast border appeared in the map of calculated spread.

As expected from a model simplification, the calculated distribution was more uniform than the observed distribution. Indeed, the incidence of disease in the two rows of 28 plots along the wind-exposed southwest border was somewhat variable (11–90%) but higher on average than elsewhere in the field. The calculated pattern of spread, although reproducing the average disease incidence, underestimated this variation, as calculated values ranged from 51 to 70% in these two rows. The underestimation of the variability was noticed in all fields but was most clearly encountered where the observed pattern of spread was highly variable. In such cases, variability is probably partly due to estimation of disease incidence based on small plots of 25 plants rather than on those of 100 plants used previously.

Kriging was used effectively to follow the evolution of the spatial pattern of spread with time. On the basis of observable symptom expression, disease incidence in field 2 was calculated for arrays

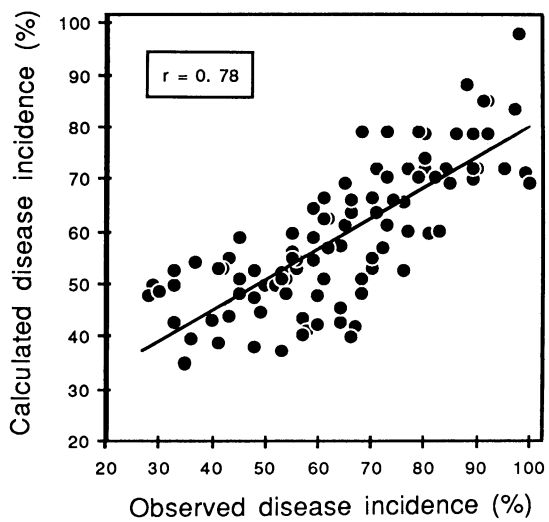

Fig. 3. Observed and calculated disease incidence in field 1, 7 mo after planting.

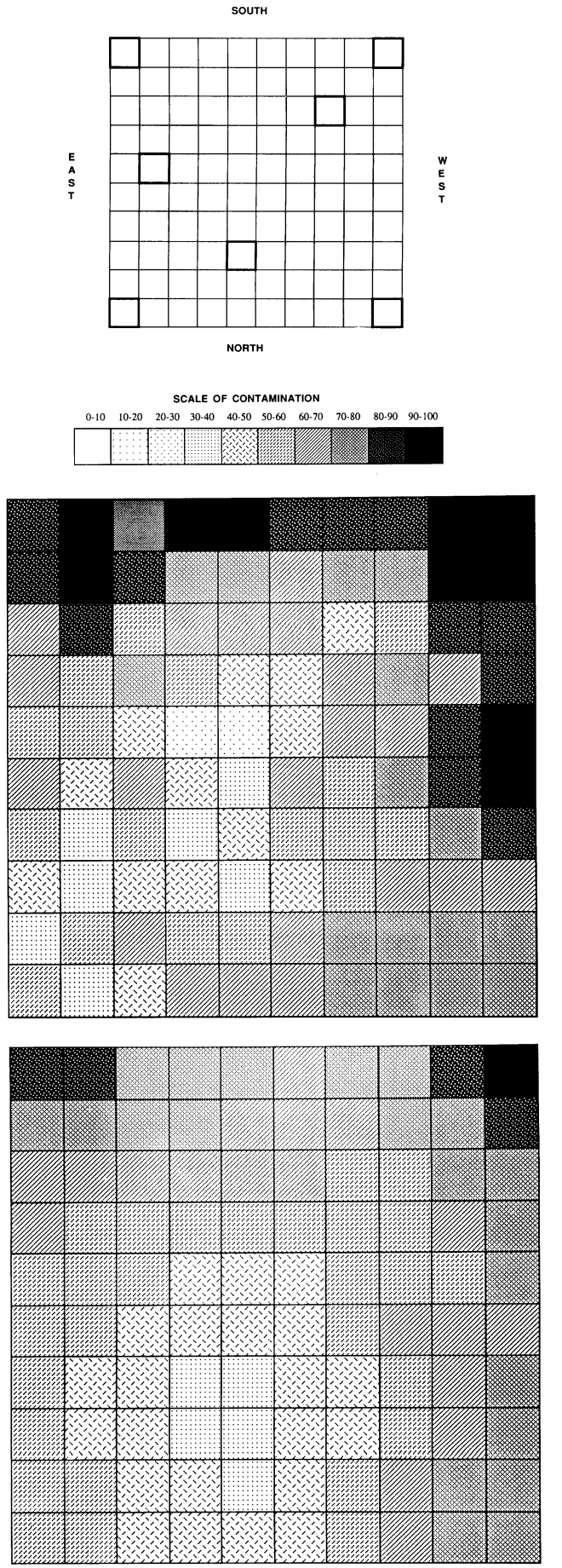

Fig. 4. Distribution of samples and orientation of field (top) and observed (middle) and calculated (bottom) mappings of disease in field 1, 7 mo after planting.

of 100 plants (by combining four adjoining subplots of 25 plants). Disease incidence was very low up to the sixth month and then increased rapidly. The level of infection was greater in the two rows of plots along the wind-exposed southwest border, whereas infection remained below 20% elsewhere. Between the sixth and eighth months, a large increase in infection occurred throughout the field, although disease incidence in the southwest wind-exposed borders remained higher than in the other borders and at the center (Fig. 7). From a sample of five blocks of the 49 plots (~10%), which comprised the center and the four corner plots, kriging clearly reproduced the main features of the spatial pattern of spread from the sixth to the eighth months (Fig. 7). However, at the sixth month, the disease pattern did not show a pronounced structure and the correlation coefficient between observed and calculated mapping was lower than those at the seventh and eighth months, when disease structure was more pronounced ($r = 0.57$, $r = 0.76$, $r = 0.76$ [df = 47] at 6, 7, and 8 mo after planting, respectively).

The observed incidence of disease in field 3 (Fig. 8, middle) was more irregular than that in field 1. No obvious disease gradient was apparent along the southwest-northeast diagonal and, unlike in field 1, high disease incidence was observed along the east border. In addition, the higher incidence was observed along the internal paths. However, a comparison of the calculated incidence of disease (Fig. 8, bottom), based on a 7% sample, with the observed incidence (Fig. 8, middle) showed that the main features of disease distribution were reproduced. Highest disease incidence was found in the southwest blocks, with high disease incidence on the southern, western, and to a lesser extent the eastern borders and lower disease incidence in the center of the field and along the northern border. If field 3 was considered to be one continuous field of 4.0 ha and the semivariogram was calculated from a 7% sample including corner plots, the correlation coefficient between calculated and observed mappings was 0.64 (df = 398). If field 3 was treated as four separate fields of 1.0 ha, the correlation coefficients were, respectively, 0.38, 0.52, 0.68, and 0.73 (df = 98). Although significant, the first two correlations were relatively low. This was likely due to the effects of the internal paths, which modify the pattern of spread, probably as a result of wind

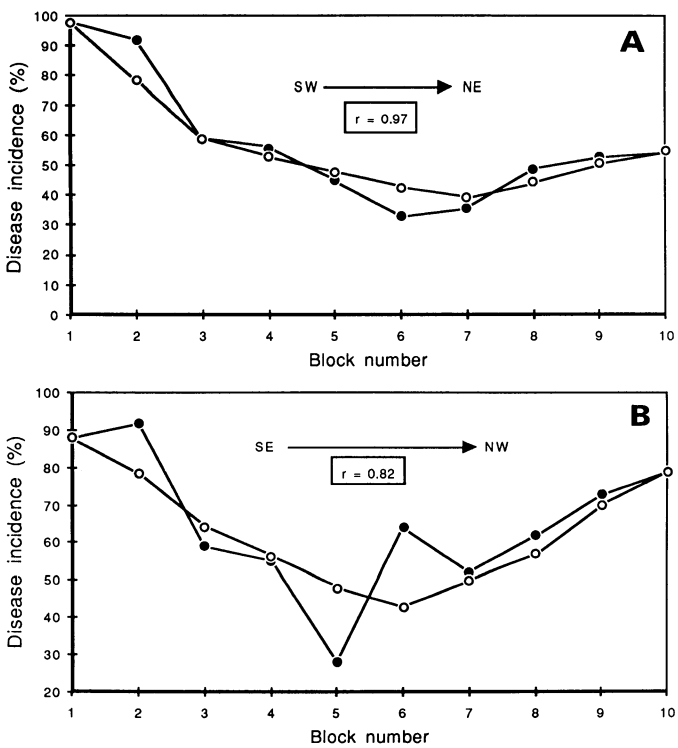

Fig. 5. Observed (●) and calculated (○) gradients of disease incidence along the southwest-northeast axis (**A**) and along the southeast-northwest axis (**B**) in field 1, 7 mo after planting.

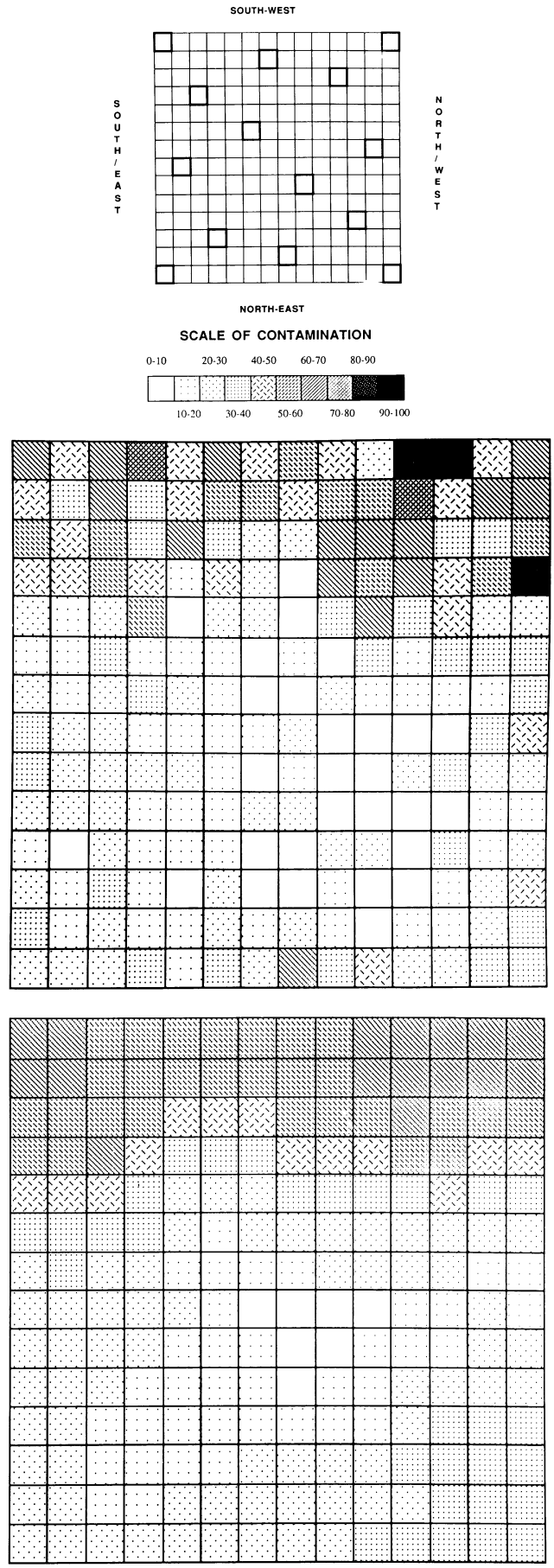

Fig. 6. Distribution of samples and orientation of field (top) and observed (middle) and calculated (bottom) mappings of disease in field 2, 6.5 mo after planting, where disease incidence is calculated for subplots of 25 plants.

turbulence and the tendency for whiteflies to settle preferentially along these paths (5). Kriging reproduced the pattern of spread more realistically when the internal border plots were excluded ($r = 0.71$, df = 322).

As the accuracy of disease estimates depends on the number and distribution of the samples taken, better estimates of spread could be obtained in field 3 with more intensive sampling. Table 2 shows the correlation between the observed and the calculated patterns of spread with different numbers of samples and window sizes. Closer correlations were obtained with larger sample sizes, provided that the size of the window was correspondingly reduced to retain a few points within the window. For instance, correlation between observed and calculated patterns could be as high as 0.78 if a 25% sample intensity were applied with a window size of 3.

DISCUSSION

Geostatistics is used in geology and pedology in diverse ways to analyze soil variation and soil genesis, to optimize sampling patterns, or for mapping properties by interpolating values from samples of limited size (21). Trangmar et al (21) predicted that geostatistics could be applied beneficially to analyze pest and disease attacks in crops. Our work has presented various ways by which geostatistics could be used to analyze the spatial distribution of a plant disease. The semivariogram (oriented and nonoriented) is a powerful tool with which to analyze the structure of spatial patterns of spread. From the shape of the semivariogram and its parameters, one can infer whether the pattern of spread

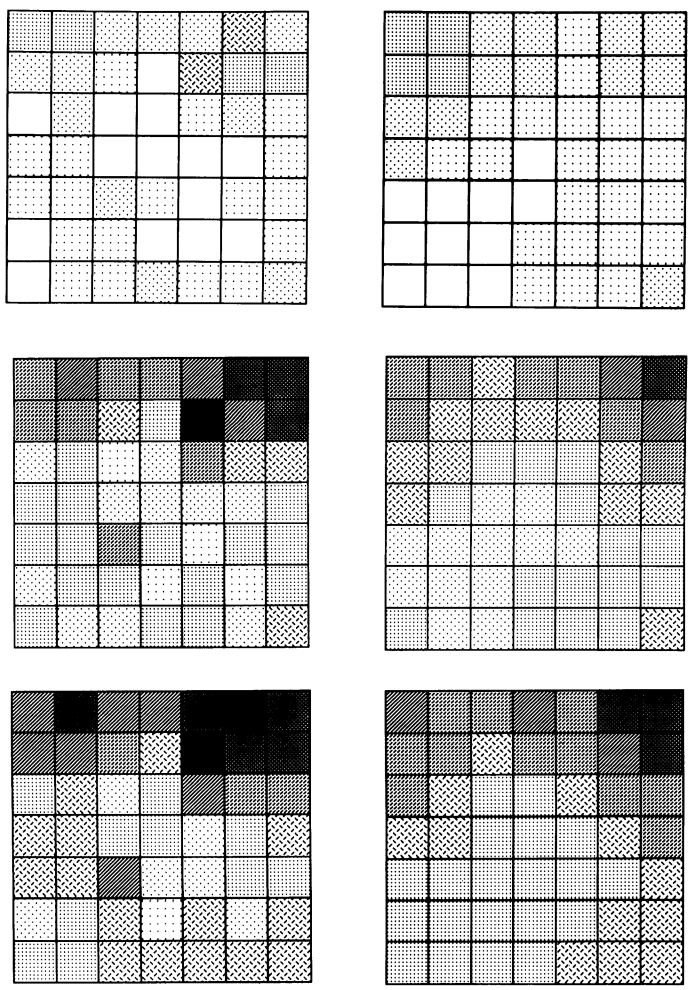

Fig. 7. Observed (left) and calculated (right) mappings of disease in field 2 at 6 (top), 7 (middle), and 8 (bottom) mo after planting, where disease incidence is calculated for plots of 100 plants. The five sample plots consisted of the four corner plots and the center plot.

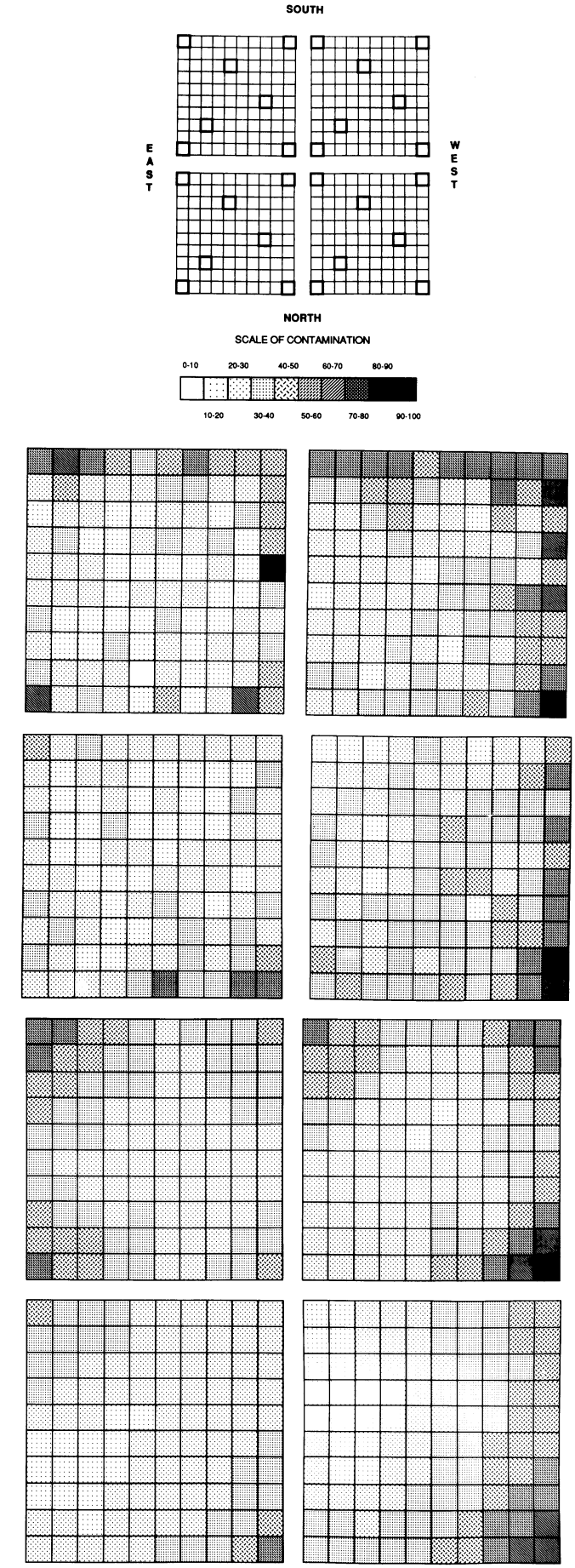

Fig. 8. Distribution of samples and orientation of field (top) and observed (middle) and calculated (bottom) mappings of disease in field 3, 4 mo after planting.

TABLE 2. Correlation coefficients between calculated and observed mappings for different sample sizes and different window sizes for field 3

Percent of sampling	Window size								
	3	5	7	9	11	13	15	17	19
7	*[a]	*	0.64	0.59	0.50	0.01	0.01	0.04	0.11
14	*	0.65	0.66	0.62	0	0.15	0.07	0.02	0.01
25	0.78	0.74	0.02	0.04					

[a] Asterisk indicates an impossible combination, as a minimum of two points is required.

is random or spatially dependent. In addition, the structural and random components can be assessed. The range of spatial dependence can be measured and the direction of maximum variation determined. With ACMV spread, the semivariograms of all the fields studied were close to linear, whereas the nugget variance was limited. This is characteristic of a strongly spatially dependent structure with limited random variation. Moreover, study of directional semivariograms indicated a strong anisotropy of the variable, with no modellike semivariogram in the northeast direction. This reflected the fact that the main direction of variation was along the southwest axis.

Some of the features of the spread of ACMV were readily apparent from direct observation, and geostatistics only confirms, refines, and quantifies the analysis. However, defining the degree of spatial dependency by deriving quantifiable parameters provides opportunities for comparative studies of spatial variability between fields and surveys. For instance, semivariograms can be used to examine how spatial patterns evolve with time—as was done for field 2, in which semivariograms were assessed at increasing times after planting—and can provide insights into the evolutionary process that led to the state of the system. Alternatively, comparisons between semivariograms can provide additional information on the mode of spread. For example, in ACMV-infected fields the semivariance increased continuously and linearly without a range, which indicates that the greater the separation of two samples, the greater the difference in disease incidence. Moreover, the semivariogram was fitted to the linear model in both fields 1 and 3, where diseased plants were retained, and in field 2, where they were removed. Similarity of shapes of the semivariograms when internal virus sources were present or removed suggested that there was little secondary spread of the disease and confirmed similar conclusions reached by other means (5). With bud rot of oil palm, semivariograms demonstrated possible secondary spread (10,11). Further features of semivariograms such as systematic deviations from the linear model provided additional information on the precise pattern of spread, such as the extent of the border effects (fields 1 and 2) and the influence of internal paths (field 3). The overall results of this analysis support the hypothesis that spread of ACMV comes mostly from sources that are outside the plantings, that its direction is associated with wind direction, and that it can be expressed over a distance and associated with wind direction over a distance exceeding the size of the field and that secondary spread is limited.

Indeed, most spatial patterns of the spread of plant diseases are not as pronounced as that of ACMV, and, as in geology, geostatistics is likely to be applied to visualize features of spatial heterogenicity that cannot be detected by direct observation, such as gradients, aggregation patterns, and ponctual anomalies. Finally, geostatistics can help to relate various biological processes that lead to the spatial pattern of spread. For instance, Chellemi et al (4) applied geostatistics to measure differences between patterns of initial inoculum in soil and patterns of diseased plants. With ACMV, progress is being made to relate the ACMV spatial pattern to whitefly vector distribution in fields (R. Lecoustre, D. Fargette, C. Fauquet, and L. Fishpool, *unpublished results*).

Several interpolation methods have been used in geology, but only kriging uses the spatial structure of the variable for estimation (7). In geology, kriging has been mostly used for isoproperty mapping, to evaluate the precision of the estimations and to define the sampling pattern in relation to the precision needed and the semivariogram (7). In this study, we assessed how effectively the kriging technique reproduced the observed patterns of spread on the basis of a sample of limited size. We found that kriging efficiently reproduced the main characteristics of ACMV distribution, including the higher incidence on the wind-exposed southwest field borders and other less obvious features. Up to 60% of the total variance was reconstructed by sampling only 7% of the total stand. Indeed, with such a sample, the precise shape of the disease gradient in field 1 was reproduced. Kriging was also applied successfully in fields differing in planting date, size, constituent plots, orientation, and mode of disease assessment (fields 2 and 3). Moreover, the technique was efficient not only in fields showing regular, expected patterns of spread but also when the pattern was heterogeneous and somewhat atypical. These results suggest that kriging is an efficient method that can be used to map different fields and that it can sometimes greatly reduce the amount of field sampling needed and yet give enough information for most epidemiological purposes. For example, in field 1, where spatial structure was highly pronounced, a 7% sampling scheme gave nearly as much information on the general pattern of spread as a more detailed sampling pattern of 25%, which provided only limited additional information. Our results show that not only the intensity but also the configuration of the sampling is important. Empirically, we found that sample patterns that take into account the corner plots allow better reconstruction of the observed spread than random sampling patterns. More generally, sampling on a grid basis is reported to be optimal and results in neighborhoods with the same number of samples (21). Establishment of such sampling grids depends on the semivariogram and the anisotropy of the variable (21).

Kriging is a robust technique, and it has been shown that minor errors in estimation of semivariogram parameters make little difference to the reliability of interpolation (21). Moreover, in ACMV-infected fields, the spherical model of the semivariogram, used instead of the linear one, gave acceptable results (R. Lecoustre, *unpublished*). This can be explained because the linear model is the beginning of the spherical model (15). However, when the spatial pattern of spread is less pronounced, only the main features of the spatial pattern are reconstructed by kriging. A high nugget variance of the semivariogram indicates a large point-to-point variation at short distances and suggests that increased sampling will often reveal more details in structure. Then, additional information will be provided by more intensive surveys, which can visualize, for example, the internal border effect in field 3. Much of the variability may occur over short distances within the sampling unit, and decreasing the sampling unit size may reveal local structures that exist at smaller scales. Actually, with the ACMV pattern of spread, successively calculating the semivariogram on size units of 25 and 100 plants in field 2 did not reveal any smaller local structure. Indeed, in extreme cases where virus-infected plants are distributed totally at random, the occurrence of disease in any one plot does not provide information on the disease incidence in any of the others. The variable is then no longer regionalized, and the semivariogram follows the "pure nugget" model. Apart from this rare situation, geostatistics and kriging can probably be used with other structures of the variable and be applied to other viral and nonviral plant diseases.

LITERATURE CITED

1. Barnett, O. W. 1986. Surveying for plant viruses. Pages 147-166 in: Plant Virus Epidemics, Monitoring, Modelling and Predicting

Outbreaks. G. D. McLean, R. G. Garret, and W. G. Ruesink, eds. Academic Press, Orlando, FL. 550 pp.
2. Bock, K. R., and Harrison, B. D. 1985. African cassava mosaic virus. AAB Description of Plant Viruses. No. 297.
3. Burgess, T. M., Webster, R., and McBratney A. M. 1981. Optimal interpolation and isarithmic mapping of soil properties. J. Soil Sci. 31:505-524.
4. Chellemi, D. O., Rohrbach, K. G., Yost, R. S., and Sonoda, R. M. 1988. Analysis of the spatial pattern of plant pathogens and diseased plants using geostatistics. Phytopathology 78:221-226.
5. Fargette, D. 1985. Epidemiologie de la Mosaïque africaine du manioc. Ph.D. thesis. University of Sciences of Montpellier, Montpellier, France. 201 pp.
6. Fargette, D., Fauquet, C., and Thouvenel, J.-C. 1985. Field studies on the spread of African cassava mosaic virus. Ann. Appl. Biol. 106:285-294.
7. Gascuel-Odoux, C. 1987. Variabilité spatiale des propriétés hydriques du sol, méthodes et résultats; cas d'une seule variable: Revue bibliographique. Agronomie 7(1):61-71.
8. Harrison, B. D. 1981. Plant virus ecology: Ingredients, interaction and environmental influences. Ann. Appl. Biol. 99:195-209.
9. Krige, D. 1966. Two dimensional weighted moving average trend surfaces for ore-evaluation. J. S. Afr. Inst. Min. Metall. 66:13-38.
10. Lecoustre, R. 1985. Contribution à l'étude de la Pourriture du Coeur de Shushufindi. Etudes statistiques épidémiologiques préliminaires. Etudes sur la régionalisation de la variable pourcentage de pertes cumulées. Rapport interne IRHO, Doc. LM/BMST No. 3.
11. Lecoustre, R. 1985. Contribution à l'étude de la Pourriture du Coeur à Palmoriente. Etudes statistiques épidémiologiques préliminaires. Etudes sur la régionalisation de la variable pourcentage de pertes cumulées. Raport interne IRHO Doc. LM/BMST No. 3 bis.
12. Lecoustre, R., and de Reffye, P. 1986. The regionalized variable theory: Possible applications to agronomical research, in particular to oil palm and coconut, with respect to epidemiology. Oleagineux 41(12):541-548.
13. Marbeau, J. P. 1976. Géostatistique forestière, état actuel et développements nouveaux pour l'aménagement de la forêt tropicale. Thèse de doctorat, Ecole des mines, Paris.
14. Matheron, G. 1963. Principles of geostatistics. Econ. Geol. 58:1246-1266.
15. Matheron, G. 1965. Les Variables Régionalisées et Leur Estimation. Une Application de la Théorie des Fonctions Aléatoires aux Sciences de la Nature. Masson, Paris. 305 pp.
16. Narboni, P. 1979. Application de la méthode régionalisée à des faits du Gabon. Note statistique No. 18. CTFT, Nogent Sur Marne.
17. Nicot, P. C., Rouse, D. I., and Yandell, B. S. 1984. Comparison of statistical methods for studying spatial patterns of soilborne plant pathogens in the field. Phytopathology 74:1399-1402.
18. Reynolds, K. M., and Madden, L. V. 1988. Analysis of epidemics using spatio-temporal autocorrelation. Phytopathology 78:240-246.
19. Reynolds, K. M., Madden, L. V., and Ellis, M. A. 1988. Spatio-temporal analysis of epidemic development of leather rot of strawberry. Phytopathology 78:246-252.
20. Thresh, J. M. 1976. Gradients of plant virus diseases. Ann. Appl. Biol. 82:381-406.
21. Trangmar, B. B., Yost, R. S., and Uehara, G. 1985. Application of geostatistics to spatial studies of soil properties. Adv. Agron. 38:45-94.