A Geostatistical Analysis of the Spatio-Temporal Development of Downy Mildew Epidemics in Cabbage

A. Stein, C. G. Kocks, J. C. Zadoks, H. D. Frinking, M. A. Ruissen, and D. E. Myers

Associate professor of Geostatistics, research associate of Plant Pathology, professor of Ecological Plant Pathology, associate professor of Plant Pathology, and assistant professor of Plant Pathology, Agricultural University, Wageningen, the Netherlands, respectively; and sixth author: professor of Mathematics, University of Arizona, Tucson.

Partly carried out with a grant from the Netherlands Integrated Soil Research Program, PCBB, during the stay of A. Stein at the University of Arizona, Tucson.

We thank A. H. C. van Bruggen, Dept. Plant Pathol., Univ. Calif., Davis; A. C. van Eijnsbergen and G. Gort, Dept. Math. Wageningen Agric. Univ., Neth., for their assistance during preparation of the manuscript; and E. Top for collecting the data.

Accepted for publication 21 June 1994.

ABSTRACT

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

Spatial and temporal data on disease incidence in an experimental field of red cabbage in which a downy mildew epidemic developed from a point source was analyzed by a geostatistical model. A two-step prediction procedure was defined to formulate a linear predictor for space and time variables based on a spatial semivariogram that changes in time. The model accommodated any pattern of spatial variables, including patterns of healthy and diseased plants. The spatial transposition of the epidemic focus was determined with high precision. Time of initiation of the epidemic was determined with lower precision due to extrapolation in time. The results were compared with data from three other experimental plots: one red cabbage plot from the previous year and two from the same year. For epidemics in white cabbage plots, a single-point source of the epidemic could not be identified. In fact, several small foci were present in the white cabbage plots. Thus, the procedures presented in this paper could distinguish between epidemics originating from a point source and those having more than one source. Optimal sampling plans based on a predetermined precision are proposed.

The downy mildew pathogen (*Peronospora parasitica* Pers.:Fr.) damages cabbage (*Brassica oleracea* L.), mainly attacking curds and heads. The disease can cause serious problems when cabbage seedlings are infected. *P. parasitica* is a biotrophic parasite that is dispersed by conidia during the crop-growing season. It survives unfavorable environments by oospores. Primary infections take place close to ground level and are often sheltered within the canopy. Secondary infections occur on higher leaf positions, and inoculum can be dispersed over greater distances (29). The latency period of the pathogen is approximately 7 days, depending on the temperature. The length of the period is affected by climatological conditions and is shortest at a canopy temperature of 16 C. In the present study, polycyclic development of the pathogen on cabbage was observed in time and space.

Epidemic development in space and time share several properties. First, observations at one specific time point are related to each other as well as to observations at earlier and later points. Second, spatial variation is often dynamic. For epidemics developing in experimental plots, two remarks can be added: Observations are usually abundantly available (e.g., 100 or 1,000) at each time point, and observations on multiple, related variables may be collected simultaneously. A statistical analysis of these data is desirable, in view of the amount of data and the objectives of the experiments. The sampling pattern may lack order and regularity both in space and time. Previous research showed the importance of analyzing the spatial distribution of pathogenic fungi in enhancing the understanding of their ecology and epidemiology (2,6,10,30,37). Gradient-models (8), autocorrelation (7,17, 18,21,25), two-dimensional distance class analysis (23), Morisita's index of dispersion (31), and geostatistics (3,12,15,35) increasingly have been used to deal with location-specific statistical inference. These procedures, however, are too restrictive to be generally

applicable to spatio-temporal data, because either the source has to be known or stationarity or fixed semivariograms have to be assumed. Theory to analyze spatio-temporal data has focused on the development and use of practical space-time models (1,7, 16,18,21,25,26,28) and on a theoretical space-time framework in relation to random fields (4).

This study deals with a dynamic spatial pattern of downy mildew epidemics on cabbage. A two-step procedure was developed to carry out predictions in space and time of a single variable, collected according to an arbitrary sampling pattern. Emphasis is given to a random function taking only the values 1 and 0, indicating presence and absence of the disease (the disease status of an individual plant).

Detailed data were available from field experiments carried out during 1991 in the Netherlands. In this paper, attention is focused on one red cabbage plot. The results were compared by those from two white cabbage plots from 1991 and with one red cabbage plot from a plot experiment carried out during 1990. This study concerns analysis of *P. parasitica* in space and time with five objectives: 1) modeling of the spatial pattern of disease at any point in time, 2) prediction of disease at unobserved points in time, 3) development of optimal sampling patterns for future assessments, 4) determination of the source of the initial inoculum in space and time, and 5) calculation of the expansion rate of the disease.

MATERIALS AND METHODS

Plot establishment. Downy mildew epidemics were studied in one cabbage plot in the Netherlands during 1990 and in three cabbage plots during 1991. Host plants for *P. parasitica* were *B. oleracea* cv. Vesta (red cabbage) in 1990 and in one plot in 1991 and *B. oleracea* cv. Stonar (white cabbage) in two plots in 1991. The experimental plots were situated at Lienden (51.57′ N latitude, 5.31′ E longitude). In 1990, the red cabbage plot (coded as RC90), consisted of 40 × 40 plants in a 0.5-m^2 grid. This plot was planted on 15 May, with rows oriented in a northeasterly direction. In 1991, all plots consisted of a 52 × 52 array of plants in a 0.5-m^2 grid. The plots were planted on 22 May with rows oriented in a northeasterly direction. The white cabbage plots, coded as WCa and WCb, were situated southwest and southeast from the red cabbage plot (coded as RC91), respectively. Plots were separated by four rows of field beans (0.5 m wide).

Inoculum preparation. An isolate of *P. parasitica* originally obtained from infected cabbage leaves was used in this study. Inoculum was prepared by suspending 48-h-old spores in sterile distilled water and adjusting the concentration to approximately 10^6 spores per milliliter. The inoculum was atomized onto the foliage of cabbage until runoff. Inoculated plants were enclosed in plastic boxes to maintain high humidity and free moisture for approximately 24 h and placed in a growth chamber at 16 C, with 16 h of light per day. Covers were removed, and symptoms were allowed to develop for 7 days in a growth chamber before the plants were planted in the plots. One inoculated plant was placed in the center of each plot to act as a point focus of disease within the plot, on 16 May 1990 and 29 May 1991. In 1991, these plants were removed on 3 June. The sources in the two white cabbage plots did not function. A natural infection, probably caused by the nearby red cabbage plot, appeared in the white cabbage plots.

Disease assessment. In RC90, all 1,600 plants were monitored from 20 May to 21 August. Plants were scored weekly, as either diseased or healthy. In RC91, only the central part consisting of 2,400 plants was monitored until 22 July in WCa and until 11 June in WCb. Thereafter, only the central 1,936 plants were monitored due to time restrictions. Two white cabbage plants died. WCa and WCb were not monitored on 19 June because the weather conditions were too bad. Temperature and relative humidity were registered by a thermohygrograph, wind direction and wind speed by a wind recorder, rain and leaf wetness by a leaf wetness recorder. In the following analysis, attention was focused on RC91, which was used to develop the statistical procedures, whereas RC90, WCa, and WCb were used subsequently to test the procedures.

Variability in space and time. Variables in space (S) and time (T) are associated with their positions, s and t. It is common to model the variability in space and time by means of a random array or field, $Z(s,t)$. Let the distance in S and T be composed of h_S in S and h_T in T. Let the transpose of a vector be denoted with a prime (′). Although such a field can have many different forms, it will be assumed throughout this study that it is stationary. For stationary random fields the expectation, $\mu_{S,T}$, is defined as $\mu_{S,T} = \mathrm{E}[Z(s,t)]$ and is independent of any translation vector, $h = (h_S,h_T)'$: $\mathrm{E}[Z(s + h_S,t + h_T)] = \mathrm{E}[Z(s,t)]$. Also, the covariance function, defined as $C_{S,T}(h) = \mathrm{Cov}[Z(s+h_S,t+h_T),Z(s,t)]$, depends solely on the lag vector, h, not on location and time. The intrinsic hypothesis is a somewhat weaker assumption that sets requirements only on differences between variables at different locations, $Z(s + h_S,t + h_T)$ and $Z(s,t)$, separated with distance h: $\mathrm{E}[Z(s + h_S,t + h_T) - Z(s,t)] = 0$ and $\mathrm{Var}[Z(s + h_S,t + h_T) - Z(s,t)]$ depends solely on h. In this case, the variance of $Z(s,t)$ need not exist. Under the intrinsic hypothesis, the semivariogram, $\gamma(h) = \gamma_{S,T}(h_S,h_T)$, in S and T is defined as

$$\gamma_{S,T}(h_S,h_T) = 1/2\mathrm{Var}\{[Z(s+h_S,t + h_T) - Z(s,t)]^2\}. \quad (1)$$

As has been shown previously (13), $\gamma(h) = C(0) - C(h)$. Throughout this study, the semivariogram will be used instead of the covariance function, because it is more general and is more reliable for estimates from the available data.

Particular cases of equation 1 include 1) the spatial semivariogram $\gamma_{S,t_0}(h_S,0)$, for a fixed time point, t_0, depending only on the Euclidean distance, h_S, between two locations in space and 2) the temporal semivariogram $\gamma_{s_0,T}(0,h_T)$, for any fixed location in space, s_0, depending only on the distance, h_T, in time:

$$\gamma_{S,t_0}(h_S,0) = 1/2\mathrm{E}\{[Z(s + h_S,t_0) - Z(s,t_0)]^2\} \quad (2)$$

and

$$\gamma_{s_0,T}(0,h_T) = 1/2\mathrm{E}\{[Z(s_0,t + h_T) - Z(s_0,t)]^2\}. \quad (3)$$

The spatial semivariogram $\gamma_{S,t_0}(h_S,0)$ depends on the distance vector in space. Different semivariograms for different directions may exist. In this case the spatial random function is called anisotropic. The alternative to anisotropy is isotropy, implying that $\gamma(h)$ depends on the length $r = |h|$ of h only, and not on its direction. If the temporal semivariograms $\gamma_{s_0,T}(0,h_T)$ are independent of the location, s_0, an average temporal semivariogram, $\gamma_T(h_T)$, reducing to $\gamma_T(r_T)$ in case of isotropy, is obtained using the semivariograms for all locations as spatial replicates and taking the expectation in equation 3 over the whole spatial domain.

In practice, the semivariogram is unknown and must be estimated from observations. Let observation times be denoted by $t_j \in T, j = 1,...,k$ and the spatial observation locations at t_j by $s_{ij} \in S, i = 1,...,N_j$. The observations, denoted by $z(s_{ij},t_j)$, are considered realizations of a random field, $Z(s_{ij},t_j)$ (19,20). In a two-dimensional space, S, $s_{ij} = (s_{1ij},s_{2ij})$, where s_{1ij} is the first coordinate and s_{2ij} is the second coordinate along orthogonal axes with respect to an arbitrary origin. The restriction to a two-dimensional area is used only for this particular study. The sample semivariogram as a function of r is computed by:

$$\hat{\gamma}(r) = \frac{1}{2N(r)} \sum_{j=1}^{N(h_T)} \sum_{i=1}^{N_i(h_S)} [z(s_{ij} + h_S,t_j + h_T) - z(s_{ij},t_j)]^2, \quad (4)$$

where $z(s_{ij},t_j)$ and $z(s_{ij} + h_S,t_j + h_T)$ is a pair of observations with a spatial distance at t_j approximately equal to h_S, the total number of such pairs being equal to $N_i(h_S)$, and a temporal distance approximately equal to h_T, the number of such pairs being equal to $N(h_T)$. The total number of pairs with separation distance equal to r equals $N(r)$.

Using the values obtained, a semivariogram model, $\gamma(r)$, may be fitted to facilitate the interpretation of patterns and to carry out predictions (22). The semivariogram models are characterized by two parameters: the sill, A_0, defined as the limiting value of the semivariogram as distance h increases, and the range parameter, b, which is associated with the distance h at which the sill is reached. The sill, and hence the range, do not necessarily exist. If the sill exists, it is equal to the variance (14). When the range parameter, b, approaches zero, the semivariogram reduces to a constant, c_0, and is termed a pure nugget effect; the sill, A_0, is then equivalent to c_0. In this case, the data are independent, and the nugget effect, c_0, appears as the (extrapolated) semivariogram value at lag zero and measures nonspatial variation such as assessment error, operator bias, and the spatial variation that occurs at very short distances (i.e., for lag lengths smaller than the smallest available distance between observations). Estimates of the parameters c_0, A_0, and b, may be obtained with any weighted nonlinear regression procedure or by means of a restricted maximum likelihood procedure (33).

A permissible semivariogram model needs to be conditionally nonnegative definite, $\mathrm{Var}(\lambda'Z) = -\lambda'\Gamma\lambda \geq 0$ for all λ with $\lambda'1_n = 0$, where Γ is obtained by evaluating the semivariogram model for the distances between the observation locations, 1_n is a vector of n elements, all equal to one, and Z is a vector containing the $Z(s_{ij},t_j)$. An example of a permissible semivariogram model is the exponential model $\gamma_{exp}(r)$, defined as:

$$\gamma_{exp}(r) = c_0\,[1 - \delta_K(r)] + A_0\,(1 - e^{-r/b}), \tag{5}$$

where the Kronecker δ, $\delta_K(r)$, is 1 for $r = 0$ and vanishes elsewhere.

It is challenging to model the variability in space and time by means of an anisotropic model. However, no natural distance exists in S and T. Even use of $\gamma(r) = \gamma(r_a)$ with

$$r_a = \sqrt{\left(\frac{r_x}{b_x}\right)^2 + \left(\frac{r_y}{b_y}\right)^2 + \left(\frac{r_T}{b_T}\right)^2}, \tag{6}$$

with r_x, r_y, and r_T the distances in the first and second spatial directions and in time and b_x, b_y, and b_T the ranges of influences in these directions, respectively, has as its main drawback that b_x and b_y are considered constant in time and that b_T is considered constant in space. In the current study, this will not be the case. Therefore, the parameters of the semivariograms are obtained for each time with sufficiently available observations (i.e., at least 100 observations), and hence, they are functions of time.

In this study, attention was given to disease status (presence and absence). For such data, the semivariogram was well-defined: equal to the probability that the variable changes as a function of distance, h, between observation locations (13).

Two-step space-time kriging. One of the objectives of this study is to predict the values at unvisited locations in space or points in time, using linear combinations of data values. Kriging provides an unbiased, minimum variance linear predictor based on the spatial dependence modeled by means of the semivariogram (5,32,36). The main issue to deal with is the prediction of the value of $Z_0 = Z(s_0,t_0)$ for arbitrary s_0 and t_0.

In this paper, we propose the use of two-step space-time kriging (2STK). This predictor is a linear combination of k predictors in space. In the first step, the predictors, $\hat{Z}(s_0,t_j) = \Sigma\lambda_{ij}Z(s_{ij},t_j)$, of the value of $Z(s_0,t_j)$ at $t_j, j = 1,...,k$ are determined. The predictors are linear in the observations, $Z(s_{ij},t_j)$, $s = 1,...,N_j$. The kriging weights, λ_{ij}, associated with each observation are determined such that the kriging variance, σ_j^2, defined as $\sigma_j^2 = \mathrm{Var}[\hat{Z}(s_0,t_j) - Z(s_0,t_j)]$, is minimal, yielding the kriging predictor as the best linear unbiased predictor. The vector of kriging weights, λ_j, is given by

$$\lambda_j = \Gamma^{-1}(\gamma_0 + \omega 1_n), \tag{7}$$

where Γ is the matrix containing the semivariogram values between the observation locations, γ_0 is the vector containing the semivariogram values between the observation locations and the prediction location, and $\omega = (1 - 1_n'\Gamma^{-1}\gamma_0)/(1_n'\Gamma^{-1}1_n)$. The weights decrease as the distance increases between the prediction location and the observation locations. If anisotropy is encountered, the weights decrease fastest in the direction with the shortest range.

In the second step, the value of $Z_0 = Z(s_0,t_0)$ is predicted, using a linear combination of the predictors $\hat{Z}(s_0,t_j), j = 1,...,k$. The best linear unbiased predictor, $\hat{Z}_0$, among all predictors with multiplicative weights is defined as:

$$\hat{Z}_0 = \sum_{j=1}^{k} \theta_j \hat{Z}\,(s_0,\,t_j) = \sum_{j=1}^{k} \sum_{i=1}^{N_j} \theta_j \lambda_{ij}\, Z\,(s_{ij},t_j), \tag{8}$$

in which θ is a vector of kriging weights. Evaluation of equation 8 in a general setting is possible only if an assumption concerning stationarity in the space-time domain is made. For example, if the intrinsic hypothesis is valid in the time dimension, the average temporal semivariogram could be used. In this study, the changes through time are modeled by means of a polynomial trend, and the observations at different times are considered to be independent of each other. The degree of the trend, ν, has to be lower than the number of sampling days, k, otherwise singularity is encountered. Therefore, if $k = 2$, $\nu = 1$, and $k \geq 3$, the degree of the polynomial trend was set equal to two. The vector of kriging weights, θ, for $\nu = 2$ is similar to the vector of regression weights given by

$$\theta = X_k(X_k'X_k)^{-1}x_0, \tag{9}$$

where X_k is a matrix of size $k \times 3$ with as its j^{th} row $(1,t_j,t_j^2)$, and x_0 is a vector equal to $(1,t_0,t_0^2)'$. If $\nu = 1$, the two kriging weights are equal to

$$\theta = [(t_0 - t_2)/(t_1 - t_2)];\ \theta = [(t_1 - t_0)/(t_1 - t_2)], \tag{10}$$

whereas for $\nu = 0$, the kriging weight equals one. The kriging variance, equal to the prediction error variance, is given by

$$\mathrm{var}(U - Z_0) = \sigma_0^2 + \theta_1^2 \cdot \sigma_1^2 + ... + \theta_k^2 \cdot \sigma_k^2, \tag{11}$$

where σ_j^2 is the kriging variance at $t_j (j = 1,...,k)$ and σ_0^2 is the variance at t_0. An estimate for σ_0^2 may be obtained as the sum of nugget effect and sill obtained from the semivariogram parameters estimated at t_0 by means of modeling the change of the nugget effect and the sill value on time as a low-order polynomial or spline.

The predictor $\hat{Z}_0$ explicitly uses spatial semivariograms that change over the course of time. As can be shown by standard methods, the predictor $\hat{Z}_0$ is an exact predictor (i.e., in observation locations the observed values are themselves predicted, and the kriging variance is equal to zero). In the case of disease prediction, the predicted values need to be rounded to one if they exceed 0.5 and to zero otherwise.

The 2STK predictor $\hat{Z}_0$ was compared in this study with two other interpolation procedures: first with general anisotropic space-time kriging (ASTK), using the anisotropic semivariogram (equation 6) and second temporal kriging (TK) was applied, assuming the intrinsic hypothesis to hold in time, allowing the use of the temporal semivariogram. Comparisons were made by suppressing, in turn, all observations at each time point $t_j, j = 1,...,k$ and predicting them using the other observations. The mean of the squared errors (MSE) between predictions and observations was determined, as well as the fraction of the correctly predicted fraction of disease status.

Some related problems. To determine the source of inoculum in space and time as well as the expansion rate, the variable $T_0(s) = T_0(s_1,s_2)$ is defined for each location, $s = (s_1,s_2)$, as the last observation time at which the plant at that location was healthy. The origin of the initial inoculum and the expansion rate of the epidemic in space and time are determined by fitting

an appropriate model through the observed $T_0(s)$ values. In this study, the empirically derived nonlinear regression equation

$$\hat{T}_0(s_1,s_2) = a_0 + b_0[(s_1 - \mu_1)^2 + g(s_2 - \mu_2)^2]^{\alpha} \quad (12)$$

is proposed, with unknown parameters a_0, b_0, μ_1, μ_2, g, and α. Estimates of the parameters were obtained with nonlinear regression. The spatial origin is given by the coordinate vector (μ_1,μ_2), the temporal origin by a_0, the expansion rate by the parameters b_0 and α, whereas parameter g measures any form of anisotropy in space.

Optimal sampling in space and time depends both on the intensity of spatial sampling and on the number of sampling days. In this study, sampling is termed optimal if the kriging variance at the most uncertain locations in space as averaged over the unsampled days, σ_{Ave}^2, is minimized. The σ_j^2, and thus σ_{Ave}^2, are independent of the actual observations. The most uncertain locations are those with coordinates in the center of neighboring observations. In this study, a neighborhood of size four was applied. A comparison with larger neighborhoods of sizes eight, 12, and 20 did not show substantial differences, although σ_{Ave}^2 values were somewhat smaller for those. Because the total number of calculations becomes intractably large (to find the average value for 10 sampling days within a 70-day study period requires the calculation of almost 10^{20} kriging variances), we determined an approximate optimum by minimizing σ_{Ave}^2 for 500 randomly

TABLE 1. Disease incidence and estimated spatial semivariogram parameters for the exponential model at different days

Plot Date	Day	No. of plants	Disease incidence	Nugget (c_0)[a]	Sill (A)[b]	Range[c]			Relative nugget[d]
						b	b_x	b_y	
RC91[e]									
6/5	0	2,400	0.07	0.05	0.04	12.81	9.69	17.94	0.52
6/10	5	2,400	0.20	0.09	0.13	10.33	7.33	26.37	0.41
6/18	13	2,400	0.50	0.16	0.13	12.82	9.37	28.64	0.56
6/25	20	2,400	0.63	0.17	0.14	33.91	20.87	202.65	0.56
7/1	26	2,400	0.96	0.04	0.01	2.03	...[f]	1.43	0.86
7/8	33	2,400	0.97	0.02	0.00	...	...	...	1.00
7/15	40	2,400	0.93	0.04	0.03	0.87	0.84	0.91	0.59
7/22	47	2,400	0.94	0.04	0.02	0.77	0.55	1.04	0.69
7/29	54	1,934	0.97	0.03	*[g]	0.83	...	...	1.00
8/5	61	1,934	0.69	0.18	0.03	0.96	0.42	20.29	0.85
8/12	68	1,934	0.49	0.22	0.03	4.00	1.25	6.96	0.89
RC90[h]									
5/23	0	1,600	*	*	...	...	...	...	1.00
5/30	7	1,600	0.01	0.01	0.01	19.32	9.60	227.62	0.36
6/7	15	1,600	0.02	0.02	0.02	5.58	5.00	6.52	0.44
6/11	19	1,600	0.11	0.05	0.09	7.38	6.39	9.16	0.38
6/19	27	1,600	0.57	0.16	...	*	*	*	...
6/26	34	1,600	0.89	0.06	...	*	*	*	...
7/3	41	1,600	0.99	0.00	0.01	0.02	0.24	0.02	0.00
7/10	48	1,600	1.00	0.00	*	0.02	0.03	0.02	0.00
7/17	55	1,600	1.00	0.00	*	0.02	0.03	0.02	0.00
7/24	62	1,600	1.00	0.00	*	0.02	0.05	0.01	0.00
7/31	69	1,600	0.96	0.00	0.03	0.02	0.34	0.02	*
8/7	76	1,600	0.74	0.18	0.01	2.03	1.83	1.59	0.94
8/14	83	1,600	0.47	0.23	0.02	9.91	5.12	21.35	0.92
WCa[i]									
6/5	0	2,400	0.01	0.01	*	0.00	0.03	...	0.93
6/11	6	2,399	0.01	*	*	0.04	0.04	0.03	0.91
6/26	21	1,934	0.02	0.00	0.01	0.02	0.03	0.03	0.00
7/2	27	1,934	0.19	0.13	0.02	11.92	26.90	8.40	0.84
7/9	34	1,934	0.31	0.11	0.11	10.02	8.30	12.08	0.48
7/16	41	1,934	0.26	0.11	0.08	7.37	4.30	14.73	0.55
7/23	48	1,933	0.53	0.11	0.15	4.46	3.27	5.78	0.43
7/29	54	1,933	0.70	0.13	0.09	3.73	2.97	4.47	0.57
8/6	62	1,933	0.96	0.04	0.01	1.10	1.24	0.93	0.84
8/13	69	1,933	0.94	0.05	*	0.22	1.00	0.16	0.96
WCb[j]									
6/5	0	2,400	0.01	0.00	0.01	0.33	0.36	...	1.00
6/11	6	2,399	0.03	0.00	0.03	0.03	0.02	0.03	1.00
6/26	21	1,935	0.16	0.11	...	...	...	...	1.00
7/2	27	1,935	0.73	0.01	0.11	6.99	13.41	3.41	0.84
7/9	34	1,935	0.66	0.05	0.18	4.23	7.23	2.47	0.48
7/16	41	1,933	0.67	0.07	0.17	4.51	8.49	2.31	0.55
7/23	48	1,933	0.89	0.06	0.05	3.03	3.21	2.81	0.43
7/29	54	1,933	0.83	0.04	...	0.01	*	0.01	...
8/6	62	1,933	0.83	0.13	0.01	1.77	0.93	...	0.84
8/13	69	1,933	0.73	0.18	0.02	7.38	4.51	...	0.96

[a]The value of the semivariogram near the origin.
[b]The limiting value of the semivariogram for large distances, h.
[c]The range for the semivariogram. Parameters b_x and b_y give the slope of the linear model in the x and y directions.
[d]Relative nugget is the ratio between c_0 and A.
[e]RC91, downy mildew in red cabbage during 1991.
[f]No data available.
[g]* = <0.01.
[h]RC90, downy mildew in red cabbage during 1990.
[i]WCa, downy mildew in white cabbage during 1991, plot a.
[j]WCb, downy mildew in white cabbage during 1991, plot b.

positioned location × times for the total number of time points equal to $k = 3,...,15$. First, from a set of 500 randomly selected triplets the three sampling days with lowest σ_{Ave}^2 were determined. Next, the minimum σ_{Ave}^2 value was determined in a set of 500 randomly selected groups of four sampling days, etc. The procedure was finished if σ_{Ave}^2 was reduced to a value that was acceptable for the practical purpose.

RESULTS

Incidence and spatial variability in time. On the first observation date, 5 June, the red cabbage crop was already infected (disease incidence = 0.07) (Table 1). In view of the latency period of approximately 7 days, this infection was probably caused by an early infection in the seedbed. During the next 4 wk, disease incidence increased to a level around 0.95 (29 July), followed by a gradual decrease to 0.5 on 12 August. The patterns of infection in RC91 at different time points are shown in Figure 1. Disease incidence was high near the source, and declined with distance from the source. The spatial pattern of diseased plants was not stationary over time. Estimated parameters of the spatial semivariograms are shown in Table 1. The spatial semivariograms obtained for the period from 5 to 25 June indicated that an exponential model adequately fitted the semivariogram values ($R^2 > 0.95$ and $df = 23$; Fig. 2).

The small relative nugget for this period indicated strong spatial dependence with limited random variation. The nugget effect, c_0, was high on 18 and 25 June, associated with high sill values on 10, 18, and 25 June and a high value of the range on 25 June. This was the result of favorable weather conditions during June (10–15 C, rain, and high relative humidity in combination with long periods of leaf wetness), which stimulated a rapid disease spread resulting in increased spatial dependence.

Direction from the source had a significant effect on disease incidence, as confirmed by the semivariograms that were highly anisotropic, with an anisotropy ratio (b_y/b_x) from 1.85 to nearly 10 at 25 June (Table 1, ranges b_x and b_y). These data showed that disease spread mainly along the southwest-northeast diagonal. The prevailing wind directions during June were west and southwest. Spatial variability dropped on 1 July and remained low during the five following weeks, because the plot was nearly completely diseased. The semivariograms were flat, showing a pure nugget effect and absence of spatial dependency (Fig. 2). On 15 July, a well-structured model with a small range (four plants) was observed and disappeared on 22 July. On 5 and 12 August, the spatial variability increased. Distribution of healthy plants was random on 5 August, because defoliation by loss of older leaves led to a decline in disease incidence during the last weeks of the experiment, and new infections did not occur because the weather was often dry and warm (frequently above 30 C). Spatial dependence returned on 12 August, showing a clustering of recovered plants. The relative nugget effect (Table 1) was above 0.4 during all weeks, indicating substantial nonspatial variation compared to the spatial variation.

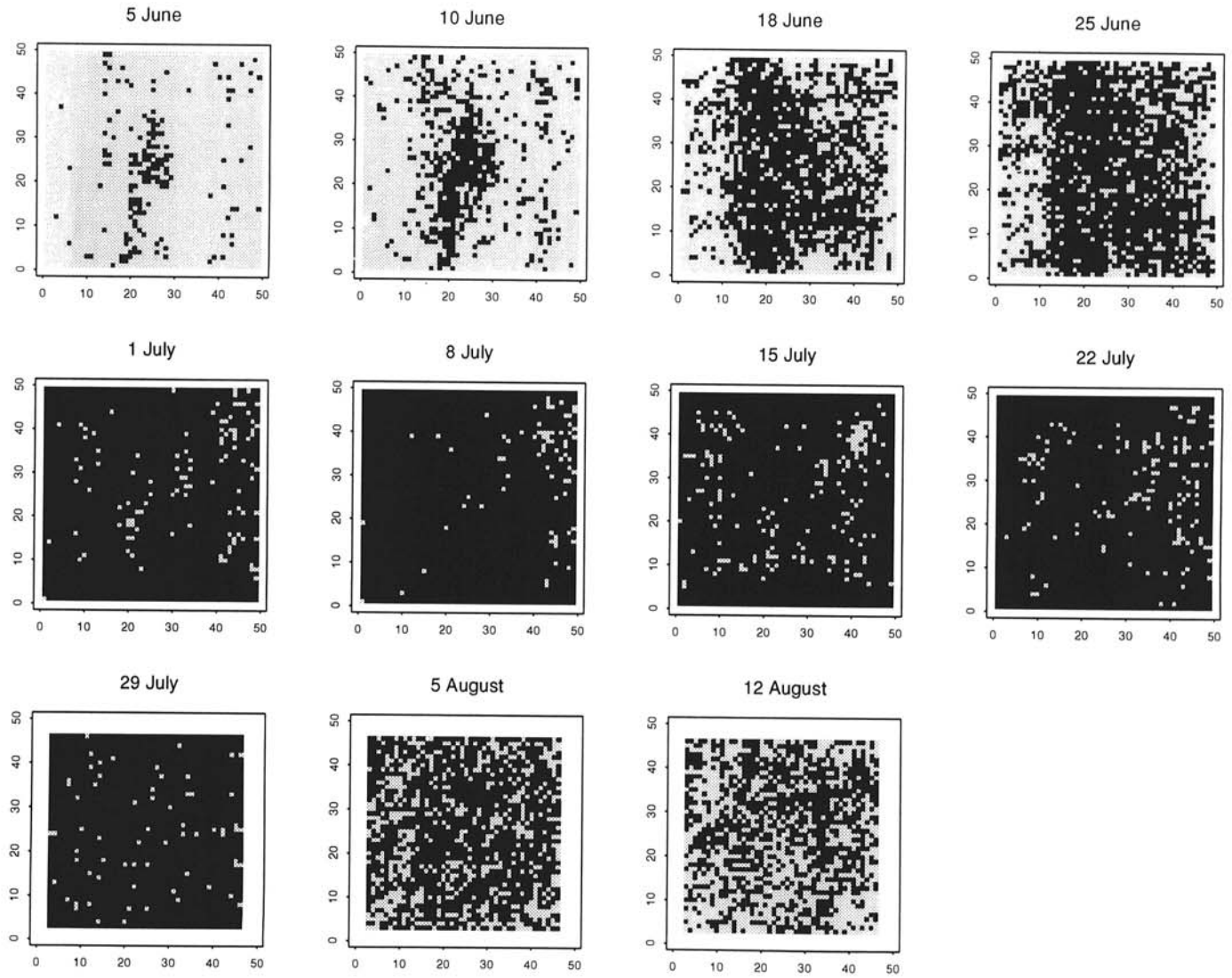

Fig. 1. The positions of the diseased and healthy plants at different points in time for downy mildew in red cabbage during 1991.

Differences (and their squares) between observations increased as a function of the time interval. The temporal semivariogram, $\gamma_T(r_T)$, is shown in Table 2. The temporal variability increased until the time interval was 7 wk. For larger time intervals, the temporal variability decreased as a result of decrease of disease incidence. The temporal semivariogram values were high (Table 2, $\gamma_T(r_T)$ values) compared to those of the spatial semivariograms (Table 1, sill values) caused by larger changes in time than at any particular day in space. The anisotropic space-time semivariogram confirmed this and showed ranges of influence in the x and y direction and in time equal to $b_x = 125$, $b_y = 137$, and $b_T = 4.77$, respectively.

Space-time kriging. The MSE for the predictions obtained with TK, ASTK, and 2STK at different days are shown in Table 3. The infection was not homogeneous throughout the plot, and the wind-exposed south and southwest borders had a lower disease incidence than other borders. This anisotropy was considered in space-time kriging. The obtained pattern, although reproducing the average disease incidence, underestimated the random variation, as could be expected from smoothing by the 2STK interpolator. In such cases, the variability was due partly to estimation of relatively rare cases of either healthy or diseased individuals. This underestimation was observed for all kriging procedures (Table 3).

The ASTK and 2STK procedures showed the lowest MSE values; the MSE value for TK was much higher. The TK procedure seemed to be less suitable than the ASTK and 2STK procedures, because the predicted fraction of infected plants was far from the actual values on 1 and 29 July and 5 August. On 25 June and later, space-time kriging allowed a good reconstruction of the observed mapping. The predictions of disease incidence obtained with the ASTK and 2STK procedures were close to the actual disease incidence, corresponding to low MSE values, because most plants were infected. Prediction of the disease incidence on 25 June, when the semivariogram had a large range and the variability was much higher than 1 wk later, and prediction on 5 and 12 August resulted in larger MSE values. On most time points, especially during the period of near-total infection, the MSE values were lowest for ASTK. This anisotropy was considered in the space-time kriging that resulted in low MSE values during the period of epidemic growth and decline (the spatial pattern was highly anisotropic). Kriging with ASTK underestimated disease incidence, especially on 18 June, whereas 2STK performed worst on 10 June (Table 4). The fraction of correctly predicted values for healthy plants was below 0.4 during the period after 25 June. Higher values were obtained for these fractions with 2STK, showing that this procedure performed better when predicting a rare event (i.e., that a plant is healthy during a period with incidence above 0.9). With disease incidence below 0.9, the fraction of correctly predicted values for infected plants was above 0.9, expected for 10 June.

Initial focus and expansion rate. A graphic display of the expansion rate is given in Figure 3. Estimation of the model parameters for $T_0(s)$ resulted in the most likely spatial origin of the disease (μ_1,μ_2) with coordinates $(25.31 \pm 0.25, 24.76 \pm 0.60)$. This reference point was not significantly different from the actual inoculation

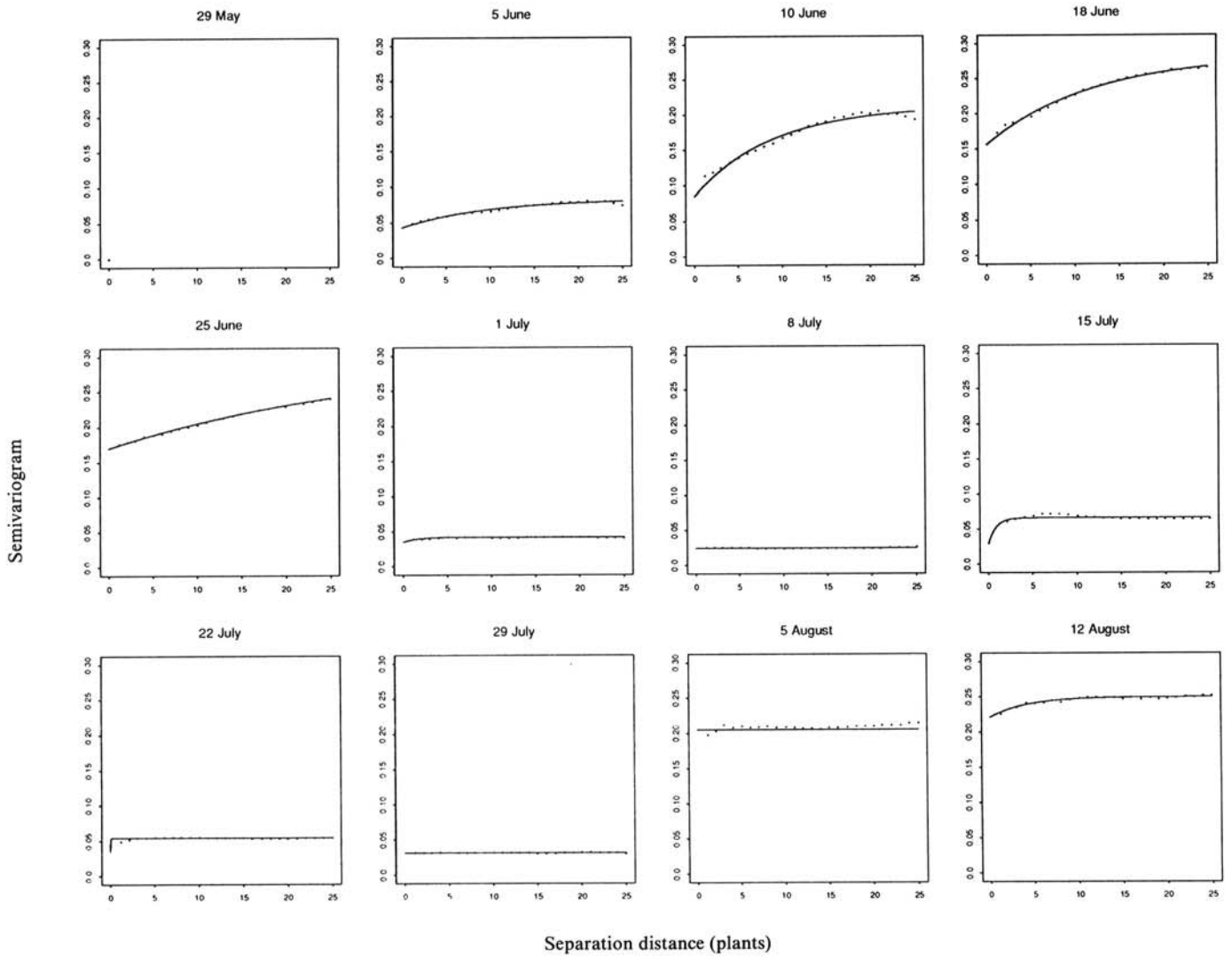

Fig. 2. The spatial semivariograms for downy mildew in red cabbage during 1991 at different points in time. Exponential semivariogram models show a clear spatial structure until 25 June and are close to a pure nugget effect afterward, with spatial structure returning at 12 August.

point (25,25), which was contained within the 95% confidence limits. The differences in the standard deviations were due to the anisotropy. The temporal origin (parameter a_0) was estimated to be -1.7 ± 1.7 days, implying that the symptom expression started after 3 June. Considering the latency period of 7 days, the epidemic started approximately 9 days before 5 June (on 27 May). This time point was 2 days before artificial inoculation (29 May) and 5 days after planting (22 May). The estimate used the last observation date at which the plant was healthy. Changing to the first observation date at which a plant is diseased would postpone the temporal origin by approximately 1 wk, supporting the conclusion that the disease originated during the week following 27 May.

Parameters b_0 and α were estimated to be 1.22 ± 0.62 and 0.42 ± 0.07, respectively. Therefore, the expansion rate, measured in days from the first observation day and modeled by $T_0(s)$, followed the power model, $1.22|h|^{0.42}$, with $|h|$ as the average distance from the source. The anisotropy in space expressed by parameter g, equal to 0.32, showed that the spread in the y direction (southwest-northeast) proceeded faster than in the x direction (southeast-northwest), as confirmed by visual inspection of Figure 1.

TABLE 2. The estimated temporal semivariograms for the incidence of diseased cabbage plants in four plots

Plot Date	No. of pairs	$\gamma_T(r_T)$[a]
RC91[b]		
6/10	22,602	0.11
6/18	20,202	0.16
6/25	17,802	0.21
7/1	15,402	0.26
7/8	13,002	0.29
7/15	10,602	0.31
7/22	8,202	0.33
7/29	5,802	0.33
8/5	3,868	0.28
8/12	1,934	0.24
WCa[c]		
6/10	17,868	0.09
6/18	15,469	0.15
6/25	13,536	0.20
7/1	11,602	0.25
7/8	9,667	0.29
7/15	7,733	0.37
7/22	5,800	0.43
7/29	3,866	0.47
8/5	1,933	0.46
WCb[d]		
6/10	17,868	0.15
6/18	15,469	0.20
6/25	13,536	0.23
7/1	11,602	0.28
7/8	9,667	0.27
7/15	7,733	0.30
7/22	5,800	0.31
7/29	3,866	0.38
8/5	1,933	0.36
RC90[e]		
30/5	19,200	0.07
7/6	17,600	0.13
11/6	16,000	0.19
19/6	14,400	0.26
26/6	12,800	0.32
3/7	11,200	0.35
10/7	9,600	0.36
17/7	8,000	0.41
24/7	6,400	0.39
31/8	4,800	0.36
7/8	3,200	0.30
14/8	1,600	0.23

[a] Estimated temporal semivariogram values.
[b] RC91, downy mildew in red cabbage during 1991.
[c] WCa, downy mildew in white cabbage during 1991, plot a.
[d] WCb, downy mildew in white cabbage during 1991, plot b.
[e] RC90, downy mildew in red cabbage during 1990.

Optimal sampling. An efficient way of mapping the spatial pattern, based on a sample of limited size but able to construct the main characteristics of disease distribution, would be a useful tool. In this experiment, we studied the effect of reducing the number of observations at any observation day by increasing the sampling distance from one to seven plants, reducing the number of observations by up to 98%. The resulting 49 observations is the minimum number needed to estimate the semivariogram. The original σ_{Ave}^2 value increased from 0.367 to 0.374, showing that this reduction can be applied without seriously affecting the quality of the predictions. The same procedure was carried out for the temporal pattern, reducing the number of sampling days. The σ_{Ave}^2 value as a function of the number of sampling days, N_{days}, is given by the equation

$$\sigma_{Ave}^2 = 0.264 + 0.620/N_{days};\ R^2 = 0.85. \quad (13)$$

The required precision dictates the number of sampling days. Optimal sampling plans are given in Figure 4. If plants are to be assessed at three points in time, the first assessment should be made at the start of the sampling period, the second one on approximately day 29 (when nearly all plants became infected), and the third at day 54, the day that recovery began. A similar schedule may be followed if more observation dates are taken. Evaluating the actual schedule, we noticed that the σ_{Ave}^2 value could be reduced from 0.367 to 0.348 by more carefully allocating the 11 observations: six observations to the period of spread, three observations to the recovery period, and two to the period of complete infection. It is especially important to assess plants during the last week and the first four weeks.

Comparison between RC91 and WCa and WCb. The epidemic in WCa proceeded much slower than in RC91 (Table 1). Disease incidence was low until 11 June. Afterward, it increased and reached its maximum, equal to 0.96, on 6 August. The proportion of diseased plants for the epidemic in WCb developed to a disease incidence of 0.73 on 2 July, after which disease incidence fluctuated between 0.66 and 0.89. The highest disease incidence was reached on 23 July. It was of interest to determine how efficiently kriging reproduces the spatial pattern of spread. The pattern of spread in WCa and WCb, although showing the greatest disease incidence in the direction of the wind (southwest-northeast), differed from that observed in RC91. Likewise the spatial semivariograms were different, expressing anisotropy first with preferential direction in the x direction (southeast-southwest), later in the y direction (southwest-northeast) for WCa. For WCb, the semivariograms expressed anisotropy with preference for the x direction (southeast-northwest) (Table 1). The temporal semivariogram attained higher values than with RC91, probably caused by generally larger differences between observations on successive assessment days in white cabbage (Table 2). The predictions of disease incidence for WCa were generally better and for WCb generally worse than those for RC91 (total MSE 0.139 and 0.234 as compared to 0.184) (Table 3). The fractions of actual and predicted infected plants are shown in Table 5.

Space-time kriging underestimated disease incidence and thus the spatial pattern of the epidemic (WCa). The model correctly predicted diseased and healthy plants for RC91 during the period of nearly complete infection of the plot (around 0.9) (Table 4). For WCb, space-time kriging estimated disease incidence with high precision. Problems arose with the estimation of parameters describing the source and the expansion rate, modeled by means of $T_0(s)$. The origin in time (a_0) was determined to be 18 (both WCa and WCb), implying that symptom expression began on 23 June. The epidemic would have started on approximately 16 June. This disagreement is probably not caused by underestimation of disease incidence with space-time kriging, because for WCb disease incidence was predicted with high precision. The origin in space was difficult to obtain for both WCa and WCb, because it was evidently outside the plot. The SAS procedure NLIN (SAS Institute, Cary, NC) failed to converge for μ_2, whereas μ_1 was

48, on the border of the plot. Because the other parameters (α, b_0, and g) are associated with the coordinates of the origin, they did not converge. Modeling ${\sigma_{Ave}}^2$ as a function of the number of sampling days was successful for WCb (15) but less successful for WCa (14):

$${\sigma_{Ave}}^2 = 0.086 + 0.164/N_{days};\ R^2 = 0.47 \quad (14)$$

and

$${\sigma_{Ave}}^2 = 0.120 + 0.294/N_{days};\ R^2 = 0.99. \quad (15)$$

The modeling for WCa was less successful because the distribution of the disease was scattered and the pattern of diseased plants was rather anisotropic. The optimal sampling plan for WCa was different from the one for RC91, with a single observation at the start of the period, a single observation at the end of the period, and the other observations between days 25 and 60 (i.e., 75% of the observations within 51% of the time period).

Comparison between RC91 and RC90. The epidemic in RC90 proceeded until a maximum disease incidence was reached on 3 July (Table 1). After 24 July, the disease incidence declined. The disease progress curve was more regular than in 1991. The spatial pattern of spread for RC90 showed a clear and well-structured focus development. The long axis of the focus that developed in the direction of the prevailing wind (southwest-northeast) also was indicated with the spatial semivariograms (anisotropy in the y direction, southeast-northwest) (Table 1). The temporal semivariogram attained larger values, caused by larger differences between time points (Table 2). The predictions in 1990 (total MSE 0.115) were more precise than in 1991 (0.184) (Table 3). Space-time kriging estimated disease incidence with high precision (Table 5). The prediction of the diseased and healthy plants was highly precise in 1990 and less precise in 1991.

TABLE 3. Mean squared error values for the three space-time kriging procedures: temporal (TK), anisotropic (ASTK), two-step (2STK)

Time (days)	RC91[a]			RC90[b]		WCa[c]		WCb[d]	
	TK	ASTK	2STK	Time (days)	2STK	Time (days)	2STK	Time (days)	2STK
0	0.154	0.115	0.149	0	0.006	0	0.017	0	0.032
5	0.154	0.208	0.155	7	0.012	6	0.017	6	0.033
13	0.280	0.309	0.280	15	0.084	NA[e]	NA	NA	NA
20	0.281	0.356	0.362	19	0.271	21	0.110	21	0.485
26	0.308	0.063	0.071	27	0.383	27	0.195	28	0.195
33	0.053	0.031	0.042	34	0.105	34	0.174	35	0.147
40	0.078	0.070	0.074	41	0.008	41	0.192	41	0.158
47	0.071	0.057	0.059	48	0.003	48	0.284	49	0.227
54	0.150	0.039	0.052	55	0.001	54	0.293	54	0.530
61	0.324	0.319	0.407	62	0.004	62	0.078	64	0.285
68	0.445	0.467	0.442	69	0.056	69	0.092	71	0.349
				76	0.243				
				83	0.325				
Total	0.209	0.180	0.184		0.115		0.139		0.234

[a]RC91, downy mildew in red cabbage during 1991.
[b]RC90, downy mildew in red cabbage during 1990.
[c]WCa, downy mildew in white cabbage during 1991, plot a.
[d]WCb, downy mildew in white cabbage during 1991, plot b.
[e]Not available.

TABLE 4. Measured and predicted fraction of diseased plants and predicted fraction of actual diseased and healthy plants by means of three prediction procedures—temporal kriging (TK), anisotropic space-time kriging (ASTK), and two-step space-time kriging (2STK)—for epidemic of downy mildew in red cabbage 1991 (RC91)

Time (day)	Fraction disease				Fraction correctly predicted			
	Actual	TK	ASTK	2STK		TK	ASTK	2STK
0	0.07	0.20	0.04	0.05	inf[a]	0.88	0.56	0.78
					n-i	0.84	0.91	0.86
5	0.20	0.07	0.12	0.06	inf	0.28	0.60	0.28
					n-i	0.99	0.84	0.99
13	0.50	0.62	0.24	0.42	inf	0.84	0.48	0.84
					n-i	0.60	0.90	0.60
20	0.63	0.52	0.62	0.60	inf	0.69	0.99	0.97
					n-i	0.76	0.07	0.09
26	0.96	0.71	0.93	0.92	inf	0.71	0.98	0.96
					n-i	0.36	0.10	0.22
33	0.97	0.96	0.97	0.95	inf	0.97	1.00	0.98
					n-i	0.33	0.08	0.18
40	0.93	0.98	0.93	0.92	inf	0.98	1.00	0.99
					n-i	0.09	0.01	0.08
47	0.94	0.97	0.94	0.94	inf	0.98	1.00	0.99
					n-i	0.11	0.00	0.07
54	0.97	0.86	0.96	0.94	inf	0.87	0.99	0.97
					n-i	0.30	0.03	0.17
61	0.69	0.93	0.66	0.49	inf	0.94	0.97	0.71
					n-i	0.10	0.06	0.33
68	0.49	0.69	0.44	0.35	inf	0.75	0.92	0.73
					n-i	0.37	0.17	0.40

[a]inf: based on number of visibly diseased plants; n-i: based on number of healthy, i.e, not visibly diseased, plants.

Estimation of the model parameters for $T_0(s)$ resulted in the most likely spatial origin of the disease (μ_1,μ_2) with coordinates (14.95 $\pm$ 0.40,13.74 $\pm$ 0.56), which differed significantly, as indicated by the values of the standard deviations, from the intended source (20,20). The temporal origin (parameter a_0) was estimated to be 14.15 $\pm$ 0.91 day, meaning that the first symptoms were estimated to appear on 29 May. Taking the latency period of 7 days into account, the epidemic would have started on 22 May (7 days after artificial inoculation), a good estimate. Parameters b_0 and α were estimated to be 0.55 $\pm$ 0.25 and 0.50 $\pm$ 0.06, respectively. Therefore, the expansion rate, measured in days from the first observation day and modeled by $T_0(s)$, followed the power model $0.55|h|^{0.50}$, with $|h|$ as the average distance from the source. The anisotropy in space (parameter g) was 0.71. Modeling ${\sigma_{Ave}}^2$ as a function of the number of sampling days was unsuccessful:

$${\sigma_{Ave}}^2 = 0.071 + 0.039/N_{days};\ R^2 = 0.45, \qquad (16)$$

as shown by the low R^2 value.

DISCUSSION

Examination of space-time analysis with RC91. Spatial variation is a property of natural populations, and it is often dynamic. Accordingly, the pattern of downy mildew changed as disease incidence increased or decreased during the course of the epidemic. The spatial pattern of the epidemic in RC91 changed in time as the disease incidence increased. The decline of disease incidence was due to natural death of infected plant parts in combination with climatologic conditions unfavorable to disease progress. At the beginning of the epidemic, a random pattern of diseased plants appeared, whereas at the end of the epidemic a random distribution of healthy plants was observed. During the progress of the epidemic, the spatial dependence between observation points increased as spatial distance decreased. There was evidence for spatial nonstationarity when the spatial dependency was determined at individual time points. This means that, at any given time, t, expected incidence level depended on location in the plot. Quantitative analysis of spatial semivariograms indicated a disease gradient during June 1991 in the direction southwest-northeast,

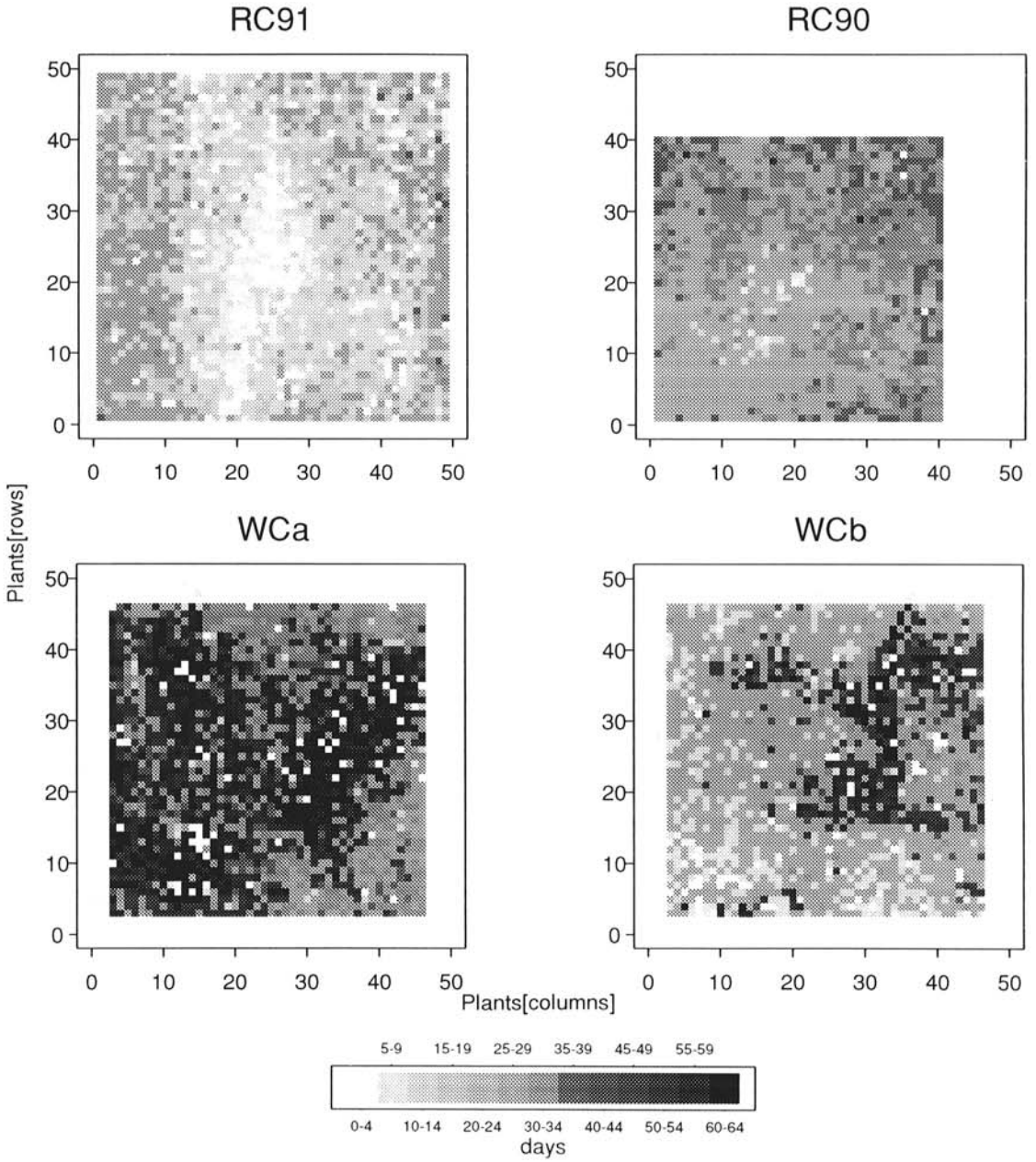

Fig. 3. The expansion rate, $T_0(s_1,s_2)$, of the epidemic is expressed for each plant as the number of days between the first observation day and the last day that the plant was observed healthy, as a function of the grid coordinates s_1 and s_2 for the four cabbage plots: red cabbage 1990 (RC90) and 1991 (RC91); white cabbage plots a (WCa) and b (WCb) 1991.

corresponding with the prevailing wind direction (southwest-west). This nonstationarity was analyzed with the semivariograms. Considering all observation times, there was strong evidence for spatial and temporal nonstationarity, indicating that expected disease status depended on the location in space and time. With these patterns, the prediction of rare events suffered from relatively large errors. This problem should be addressed in future studies. The spread of the pathogen in RC91, expressed by the function $T_0(s)$, followed the power model $1.22|h|^{0.42}$. This function was useful in determining the spatial and temporal origin of the disease. Extrapolation of this function toward the edges of the plot, however, was quite unrealistic. For such purposes, the function needs to be refined, because it should have a vertical asymptote toward larger distances between spatial points. Larger distances, between a source and healthy plants, should result in a longer time period before a healthy plant becomes infected.

Comparison among RC91, RC90, WCa, and WCb. The comparison of plot RC91 with RC90, WCa, and WCb showed the capacity of the model presented in this paper to identify the sources of disease in space and time. It showed that the $T_0(s)$ variable can distinguish between a disease pattern induced by multiple sources and a pattern induced by a point source. If infection is caused by a point source, its location can be determined with high precision, whereas the $T_0(s)$ variable does not converge if disease is caused by a multiple source.

Space-time kriging analysis was used to identify a model that would account for temporal and spatial observations. The analysis suggested that observed spatial patterns were related to temporal nonstationarity. The 2STK analysis illustrated the development of patterns over time but provided no information on the mechanisms that accounted for this development. Also, 2STK predicted disease progress, but implications of the 2STK model structure were difficult to visualize because the prediction of diseased and nondiseased plants was unsuccessful (Table 4). The 2STK predictor is attractive, because it deals with changing semivariograms over the course of time. However, the 2STK technique should be refined. It is strange that within the same year, the disease incidence was predicted with high and low precision (underestimation) for WCa and WCb, respectively. Although other predictors in space and time also may give low MSE values (Table 3), they lack the basic property of dealing with changing spatial variability. Modeling the changes in time by means of a polynomial trend could be further refined, but some assumptions concerning stationarity in the time domain are required.

At the beginning of these analyses we had hoped to be able to define an appropriate sampling scheme for spatial and temporal sampling by examining dependencies in disease incidence in space and time. Unfortunately, the specification derived from the prescribed technique was less than optimal with respect to the temporal analysis. Modeling of σ_{Ave}^2 as a function of the number of sampling days was less successful for RC90 and WCa. It appears that this approach partially failed (WCb and RC90) as a consequence of aggregation in disease incidence data, anisotropy, temporal nonstationarity, and interacting foci. Problems arose when estimating the spatial origin for WCa and WCb. An explanation might be that the actual sources were located outside the plots and that the artificial inoculations failed. The fact that the disease probably came from outside the plot could explain the change in anisotropy, because on 3 July the pathogen could not spread further in the x direction, and hence anisotropy in the y direction was observed (WCa). However, in that case, the epidemics of 1991 would all have been initiated from sources outside the plots. It is unrealistic to state that both WCa and WCb were infected by downy mildew from RC91. In that case, the disease would have spread against and perpendicular to the prevailing wind direction. We did not search for sources outside the plots. A better explanation is that the infected plants were distributed randomly over the plot (WCa and WCb) from 5 until

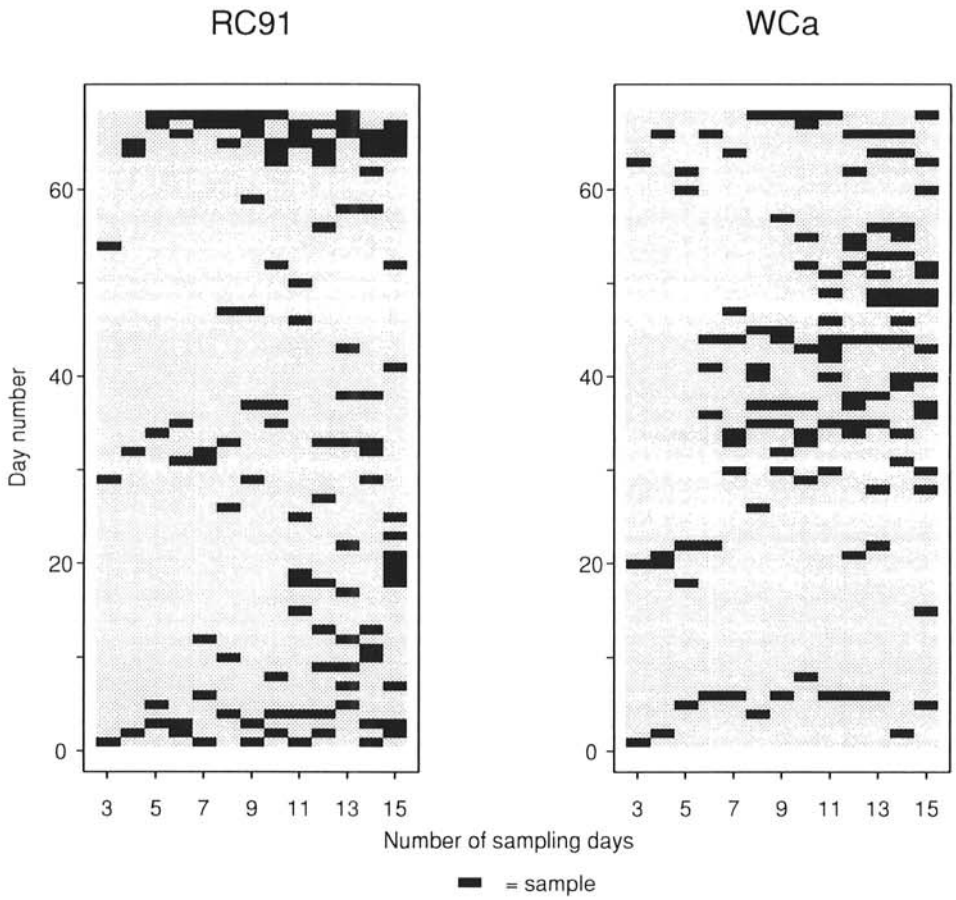

Fig. 4. Optimal sampling plans as a function of the number of observations for downy mildew in red cabbage during 1991 (RC91) and for downy mildew in white cabbage plot a (WCa) during 1991. Every black block represents an observation day.

26 June, a period during which disease incidence was below 0.02 and that some of the infected plants acted as sources of inoculum and developed as single foci. The calculations were based on the hypothesis of a single focus, and thus, problems may arise when more than one focus is active. For RC90, the extrapolated spatial and temporal origins differed from the actual origins. According to the analysis, the first symptoms could be seen on 29 May, a time point at which a clear elliptical focus had already developed. The long axis of the developing focus ran parallel to the prevailing winds (southwest-northeast) and had its own central point. Extrapolation in time of the spatial source is based on the configuration of the disease incidence values within the plot and can be erroneous. Because we are dealing with the disease status of individuals, 2STK cannot distinguish between plants with high and low severity values.

The case study presented in this paper focuses on the use of space-time data analysis for downy mildew in cabbage in four situations only. Data of more growing seasons and other pathosystems will be necessary to further evaluate the described approach. In the present study, we dealt with incidence values. Severity values also can be used but should be tested first to get more insight into the strength of the analysis. Some further refinements could be included to better describe the spatio-temporal development of the disease. For example, the epidemic process has a cause-and-effect relationship. That is, the pattern of disease at t_2 is caused by the pattern of disease at t_1 if $t_1 < t_2$, whereas the model applied in this study measures the linear relationships between the observations only at t_1 and t_2. This limitation is not considered a major drawback, because modeling and interpreting spatial variability may yield sufficient insight for many practical purposes. However, this model can be extended to include a Box-Jenkins type of study (18) in the time domain, though it will usually need more data and special assumptions.

The usefulness of the analysis is not limited to phytopathology. Many agricultural and environmental phenomena exhibit similar spatial patterns of development over the course of time. For example, the distribution of a pollutant in a homogeneous medium, the changes in groundwater levels in an aquifer, or the changing patterns on successive satellite imagery. Analyses in the space-time domain have been carried out to serve a broad range of purposes, for example to determine a rainfall network (26), to analyze piezometric readings (27), to analyze the effects of sulphate deposition (1), to assess Ireland's wind power resource (9), and to determine the nature of periodontal disease progression (34). Space-time analyses as discussed in this paper may have a future in plant disease epidemiology.

APPENDIX

Geostatistical glossary. The geostatistical glossary explains terms from the text (11,24). The interested reader is referred to Isaaks and Srivastava (11) and Olea (24) for fuller explanations. Throughout the glossary, we consider a regionalized variable, $Z(x)$, where x takes values in a spatial domain, S, and Z represents the spatially varying properties. At one single location, x_i, the regionalized variable, $Z(x_i)$, may be observed, showing a univariate distribution. An observation made at this location is denoted with $z(x_i)$. If one considers the regionalized variables at two different locations, a bivariate distribution is encountered.

TABLE 5. Measured and predicted fraction of visibly diseased plants and predicted fraction of actually diseased and healthy plants by means of the two-step space-time kriging (2STK)—of downy mildew epidemics in red cabbage 1990 (RC90) and white cabbage 1991 (WCa and WCb)

Fraction diseased

RC90			WCa			WCb		
Time	Actual	Pred[a]	Time	Actual	Pred	Time	Actual	Pred
0	0.00	0.00	0	0.01	0.00	0	0.01	0.01
7	0.00	0.01	6	0.01	0.00	6	0.03	0.03
15	0.02	0.02	NA[b]	NA	NA	NA	NA	NA
19	0.11	0.11	21	0.02	0.00	21	0.16	0.15
27	0.57	0.57	27	0.19	0.08	28	0.73	0.73
34	0.89	0.89	34	0.31	0.19	35	0.66	0.65
41	0.99	0.99	41	0.26	0.18	41	0.67	0.66
48	1.00	1.00	48	0.53	0.31	49	0.89	0.89
55	1.00	1.00	54	0.70	0.50	54	0.43	0.43
62	1.00	1.00	62	0.96	0.91	64	0.83	0.82
69	0.96	0.96	69	0.94	0.90	71	0.73	0.72
76	0.74	0.74						
83	0.47	0.47						

Fraction correctly predicted

RC90			WCa			WCb		
Time	inf[c]	n-i[d]	Time	inf	n-i	Time	inf	n-i
0	1.00	0.99	0	0.16	0.99	0	0.42	0.97
7	0.50	0.99	6	0.26	0.99	6	0.16	0.99
15	1.00	0.91	NA[b]	NA	NA	NA	NA	NA
19	0.59	0.75	21	0.13	0.91	21	0.88	0.45
27	1.00	0.11	27	0.45	0.89	28	0.78	0.86
34	1.00	0.03	34	0.63	0.91	35	0.90	0.77
41	1.00	0.33	41	0.70	0.85	41	0.92	0.70
48	1.00	1.00	48	0.58	0.86	49	0.77	0.80
55	1.00	1.00	54	0.72	0.96	54	0.87	0.17
62	1.00	0.83	62	0.95	0.20	64	0.80	0.33
69	0.97	0.19	69	0.96	0.15	71	0.79	0.28
76	0.96	0.18						
83	0.97	0.42						

[a] Predicted.
[b] Not available.
[c] inf: based on number of visibly diseased plants.
[d] n-i: based on number of healthy plants.

Translation invariance: A regionalized variable is called translation invariant if its basic properties, such as mean, variance, covariance structure, etc., are similar at different locations. Translation invariance extends well to univariate and bivariate distributions.

Translation vector: A translation vector separates two locations at which the regionalized variables are observable or have been measured.

2nd order stationarity: Second order stationarity, in addition to translation invariance of any bivariate distribution, implies that the covariance exists and depends only on the distance, h, between pairs of locations and not on their individual spatial locations. This further implies that both the variance and the mean exist and do not depend on spatial locations.

Intrinsic hypothesis: A somewhat weaker assumption than second order stationarity is summarized as the intrinsic hypothesis:
1) The expectation of $Z(x)$ exists and is independent of the location, x;
2) For all vectors, h, the increment $Z(x) - Z(x + h)$ has a finite variance, independent of x, half of which is equal to the semivariogram.

Lag vector: The distance vector, h, between two locations where the regionalized variable can be (or is) observed.

Semivariogram: For any lag vector, h, the semivariogram is defined as half the expectation of the squared difference of the random field at locations separated by that distance: $\gamma(h) = 1/2\mathrm{E}[Z(x) - Z(x + h)]^2$, where E denotes the mathematical expectation.

Range: The range of a semivariogram is the maximum distance separating points of a regionalized variable that has any significant statistical dependence. The range is the smallest semivariogram argument for which the semivariogram is either exactly equal to the sill or asymptotically close to the sill.

Sill: The limiting value for large arguments of a semivariogram.

Nugget effect: An apparent discontinuity in the experimental semivariogram near the origin.

Relative nugget: The ratio between the nugget effect and the sill value, indicating which fraction of the variability is of a nonspatial nature.

Isotropy: The characteristic of a semivariogram to depend only on the length, $|h| = \sqrt{(h_1^2 + h_2^2 + h_3^2)}$, of the separation vector, and not on its direction.

Anisotropy: The quality or state not possessing isotropy.

Conditionally nonnegative definite: A semivariogram, $\gamma(h)$, is said to be conditionally nonnegative definite if

$$-\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j \gamma(x_i - x_j) \geq 0$$

for all permissible coefficients $(a_1, a_2, \ldots, a_m)$.

Kronecker δ: The function $\delta(x)$, which is equal to one if $x = 0$ and vanishes elsewhere.

Kriging: A collection of interpolation methods, based on predicting the value of a random function at an unvisited (nonsampled) point in space and time.

LITERATURE CITED

1. Bilonick, R. A. 1985. The space-time distribution of sulphate deposition in the northeastern United States. Atmos. Environ. 19:1829-1845.
2. Campbell, C. L., and van der Gaag, D. J. 1993. Temporal and spatial dynamics of microsclerotia of *Macrophomina phaseolina* in three fields in North Carolina over four to five years. Phytopathology 83:1434-1440.
3. 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.
4. Christakos, G. 1992. Random Field Modelling in the Earth Sciences. Academic Press, San Diego, CA.
5. Cressie, N. A. C. 1991. Statistics for Spatial Data. Wiley & Sons, New York.
6. Daminicone, J. P., Snow, J. P., and Berggren, G. T. 1990. Spatial and temporal spread of soybean stem canker from an inoculum point source. Phytopathology 80:571-578.
7. Gottwald, T. R, Reynolds, K. M., Campbell, C. L., and Timmer, L. W. 1992. Spatial and spatiotemporal autocorrelation analysis of citrus canker epidemics in citrus nurseries and groves in Argentina. Phytopathology 82:843-851.
8. Gregory, P. H. 1968. Interpreting plant disease dispersal gradients. Annu. Rev. Phytopathol. 6:189-212.
9. Haslett, J., and Raftery, A. E. 1989. Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource (with discussion). Appl. Stat. 38:1-50.
10. Headrick, J. M., and Pataky, J. K. 1988. Spatial and temporal development of common rust in susceptible and partially resistant sweet corn hybrids. Phytopathology 78:227-233.
11. Isaaks, E. H., and Srivastava, R. M. 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.
12. Johnson, D. A., Alldredge, J. R., Allen, J. R., and Allwine, R. 1991. Spatial pattern of downy mildew in hop yards during severe and mild disease epidemics. Phytopathology 81:1369-1374.
13. Journel, A. G. 1986. Geostatistics: Models and tools for the earth sciences. Math. Geol. 18:119-140.
14. Journel, A. G., and Huijbregts, C. J. 1978. Mining Geostatistics. Academic Press, New York.
15. Lecoustre, R., Fargette, D., Fauquet, C., and de Reffye, P. 1989. Analysis and mapping of the spatial spread of african mosaic virus using geostatistics and the kriging technique. Phytopathology 79:913-920.
16. Loader, C., and Switzer, P. 1989. Spatial covariance estimation for monitoring data. SIMS Tech. Rep. 133. Stanford University, Department of Statistics, Stanford, CA.
17. Madden, L. V., Pirone, T. P., and Raccah, B. 1987. Analysis of spatial patterns of virus-diseased tobacco plants. Phytopathology 77:1409-1417.
18. Madden, L. V., Reynolds, K. M., Pirone, T. P., and Raccah, B. 1988. Modelling of tobacco virus epidemics as spatio-temporal autoregressive integrated moving-average processes. Phytopathology 78:1361-1366.
19. Matheron, G. 1965. Les Variables Régionalisées et Leur Estimation. Masson, Paris.
20. Matheron, G. 1973. The intrinsic random functions and their applications. Adv. Appl. Probab. 5:439-468.
21. Mihail, J. D. 1989. *Macrophomina phaseolina*: Spatio-temporal dynamics of inoculum and of disease in a highly susceptible crop. Phytopathology 79:848-855.
22. Myers, D. E. 1988. Interpolation with positive definite functions. Sci. Terre 28:251-265.
23. Nelson, S. C., Marsh, P. L., and Campbell, C. L. 1992. 2DCLASS, a two-dimensional distance class analysis software for the personal computer. Plant Dis. 76:427-432.
24. Olea, R. A. 1991. Geostatistical Glossary and Multilingual Dictionary. Oxford University Press, New York.
25. 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.
26. Rodriguez-Iturbe, I., and Mejia, J. M. 1974. The design of rainfall networks in time and space. Water Resour. Res. 10:713-718.
27. Rouhani, S., and Myers, D. E. 1990. Problems in space-time kriging of geohydrological data. Math. Geol. 22:611-623.
28. Rouhani, S., and Wackernagel, H. 1990. Multivariate geostatistical approach to space-time data analysis. Water Resour. Res. 26:585-591.
29. Royle, D. G., and Kremheller, H. T. 1981. Downy mildew of the hop. Pages 359-419 in: The Downy Mildews. D. M. Spencer, ed. Academic Press, New York.

30. Savary, S. 1986. The effects of age of the groundnut crop on the development of primary gradients of *Puccinia arachidis* foci. Neth. J. Plant Pathol. 93:15-24.
31. Shuh, W., Frederiksen, R. A., and Jeger, M. J. 1986. Analysis of spatial patterns in sorghum downy mildew with Morisita's index of dispersion. Phytopathology 76:446-450.
32. Stein, A., and Corsten, L. C. A. 1991. Universal kriging and cokriging as a regression procedure. Biometrics 44:575-588.
33. Stein, A., Staritsky, I. G., Bouma, J., van Eijnsbergen, A. C., and Bregt, A. K. 1991. Simulation of moisture deficits and interpolation by universal cokriging. Water Resour. Res. 27:1963-1973.
34. Sterne, J. A. C., Kingman, K., and Löe, H. 1992. Assessing the nature of periodontal disease progression—An application of covariance structure estimation. Appl. Stat. 41:539-552.
35. Van de Lande, H. L. 1993. Studies on the epidemiology of spear rot in oil palms (*Elaeis guineensis* Jacq.) in Suriname. Ph.D thesis. Wageningen Agricultural University, Wageningen, the Netherlands.
36. Webster, R. 1985. Quantitative spatial analysis of soil in the field. Pages 1-70 in: Advances in Soil Science, vol. 3. B. A. Stewart, ed. Springer Verlag, New York.
37. Yang, X. B., and Zeng, S. M. 1992. Detecting patterns of wheat stripe rust pandemics in time and space. Phytopathology 82:571-576.