

## Analysis and Comparison of Fusiform Rust Disease Progress Curves for Five Slash Pine Families

M. M. Griggs, W. L. Nance, and R. J. Dinus

Geneticist, Plant Geneticist, and former Principal Plant Geneticist, Southern Forest Experiment Station—U.S. Dep. Agric., Forest Service, Gulfport, MS 39503. Dinus is presently Manager, Western Forest Research Center, International Paper Co., Portland, OR 97208.

Accepted for publication 17 May 1978.

### ABSTRACT

GRIGGS, M. M., W. L. NANCE, and R. J. DINUS. 1978. Analysis and comparison of fusiform rust disease progress curves for five slash pine families. *Phytopathology* 68:1631-1636.

Disease progress curves were employed to quantify genetic variation in fusiform rust (caused by *Cronartium fusiforme*) resistance among five half-sib slash pine families planted at two locations. A growth analysis procedure was adopted whereby disease progress curves for each family were smoothed by fitting polynomials in time to cumulative percent infection. Then the coefficients for each polynomial fit were subjected to univariate and multivariate analyses of

variance. Family effects were strong and statistically significant; location and location  $\times$  family effects were identifiable but not statistically significant. Three distinct categories of disease progress curves were found: resistant, intermediate, and susceptible. Methods are advanced to quantify and evaluate the importance of factors that affect disease progress in fusiform rust and other plant diseases.

*Additional key words:* *Cronartium fusiforme*, *Pinus elliotii* var. *elliotii*, epidemiology.

A useful approach to studying epiphytotics is to quantify disease progress over time using "disease progress curves" and to predict changes in such "curves" brought about by different factors. This approach has led to increased understanding and better control of many agronomic plant diseases (11). Forest pathologists have applied disease progress curve analyses to a few forest tree disease caused by autoecious pathogens—oak wilt [pathogen: *Ceratocystis fagacearum* (Bretz) Hunt] (15), annual canker of sugar maple [pathogen: *Fusarium solani* (Mart.) App. & Wr. em. Snyder & Hans.] (23), hypoxylon canker of aspen [pathogen: *Hypoxylon mammatum* (Wahl.) Miller] (13)—and to white pine blister rust which is caused by a heteroecious pathogen (*Cronartium ribicola* Fischer) (14).

Fusiform rust (caused by *Cronartium fusiforme* Hedgc. & Hunt ex Cumm.) a major disease of slash (*Pinus elliotii* var. *elliotii* Engelm.) and loblolly (*P. taeda* L.) pines, however, has not been subjected to disease progress curve analysis. Like white pine blister rust, fusiform rust has a complex life cycle: it is heteroecious and macrocyclic, with pycnial and aecial stages occurring on pine hosts. Uredial, telial, and basidiospore stages occur on the oak host (primarily *Quercus nigra* L.). In the spring, wind-disseminated basidiospores colonize succulent pine tissues. Subsequently, the fungus grows into branch and stem tissues forming spindle-shaped swellings or galls. Stem galls caused mortality as young trees are girdled and older trees are predisposed to wind breakage. Conservative estimates indicate that fusiform rust reduced the 1972 loblolly and slash pine harvest by

\$28 million (16).

Research has provided some information on the climatic conditions that favor infection (20), the edaphic factors that affect pine infection (9), the importance of various alternate hosts and their abundance (6, 8), genetic resistance in the pine hosts (7, 18, 21), and variation in virulence of the pathogen (5). The effects of such factors have not been integrated and related to disease increase over time. A first step toward integration is the statistical comparison of disease progress curves for different genetic materials in common environments or for genetically similar materials in contrasting natural or experimentally altered environments.

This report describes statistical methods for (i) fitting and comparing cumulative disease progress curves, and (ii) quantifying the effect of five open-pollinated slash pine families and two planting locations on fusiform rust disease progress curves.

### MATERIALS AND METHODS

**Plantations.**—Open-pollinated progenies from five south Mississippi slash pines were established in 1963 at Gulfport, Mississippi (Location 1), and in 1964 on Crown Zellerbach Corporation land near Bogalusa, Louisiana (Location 2). Families 1, 2, and 3 were from rust-free phenotypes and families 4 and 5 were from rust-infected phenotypes (4, 10). The Gulfport planting was a randomized block design. Each family was represented by a 30-tree plot in each of seven blocks. The Bogalusa planting consisted of five blocks, each containing a 30-tree plot for each of the five families.

**Measurements.**—The disease frequency measure used for each plot in this study was:

This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source. The American Phytopathological Society, 1978.

$$y(t) = \frac{(\text{no. trees ever infected through } t) + (\text{no. rust-killed trees through } t)}{(\text{no. living trees at } t) + (\text{no. rust-killed trees through } t)}$$

where  $t$  = time, and  $y(t)$  = disease frequency at time  $t$ . In developing this cumulative estimate, trees dying of causes other than rust were purged from the data at and after the year in which they died. Such mortality mainly occurred in the first 2 yr after planting, but was small, was not correlated to relative family resistance (4), and did not influence the outcomes of analyses.

Measurement schedules differed at the two locations. Location 1 was measured annually at the end of the first six growing seasons and at the end of the ninth, tenth, and twelfth growing seasons. Location 2 was measured annually at the end of the first four growing seasons and again at the end of the sixth, tenth, and twelfth growing seasons.

**Analyses.**—For the purpose of this study, we considered a cumulative disease progress curve as the basic unit of response (12). These curves consisted of cumulative disease frequency measurements [ $y(t)$ 's] plotted against years after planting ( $t$ ) for each experimental plot of 30 test trees. Although graphic comparison of such curves showed differences between locations and families, the proposed methods provided a more objective and quantitative approach to such comparisons.

We adopted a growth analysis procedure proposed by Rao (17) and Wishart (24) to provide for statistical comparison of entire disease progress curves. Basically, Rao's procedure consisted of (i) smoothing the curves by fitting polynomials in time to the measurements for each plot, (ii) substituting the coefficients of the polynomial fit of each plot for the raw data, and (iii) statistically comparing the coefficients as response variables.

This method had definite advantages because (i) statistical comparison was possible whereas it may not have been in other methods, and (ii) the disease progress curves did not have to be based on the same measurement schedule. By first fitting smooth polynomials in time to the curves, and then using the estimated coefficients as data, statistical comparison was facilitated. Two curves that "look alike," even though they are based on different measurement schedules, yield similar coefficients in the polynomial fit.

**Polynomial curve fitting.**—A considerable amount of curve-fitting work was necessary in this approach, since each experimental plot (60 in all) had to be fitted with polynomials in time. The degree of polynomial must be high enough to preserve most of the original information, yet achieve some parsimony. For example, a disease progress curve with 10 measurement points may be fitted exactly with a ninth-degree polynomial. Substitution of the 10 coefficients, including the intercept, for the 10 original measurements results in no loss of information (17), but there is no gain in parsimony from such a fit. Generally, a polynomial of much lower degree provides an adequate fit (as measured by  $R^2$ ), and increasing the degree above this adequate minimum does not necessarily give a better fit. Beaton and Tukey (1) provide excellent

guidance in polynomial curve-fitting procedures.

Generally, the form of the fitted polynomial for each plot appeared as:

$$y(t) = a_0 + a_1t + a_2t^2 + a_3t^3 + \dots + a_pt^p$$

where  $y(t)$  was the predicted disease frequency (cumulative) at time  $t$ , and the  $a$ 's were the polynomial coefficients estimated by least squares. We assumed the deviations from prediction (errors) were distributed normally with mean zero and common variance. The degree of each polynomial ( $p$ ) must be the same for each plot; otherwise the number of coefficients would not be the same from plot to plot—considerably complicating analysis.

Depending upon curve shape and prior knowledge of the biological system, polynomial form may be modified by deleting some lower-order terms. For example,  $a_0$  represents the amount of disease present in the plot at planting—zero because only rust-free seedlings were planted. We therefore assumed  $a_0$  to be zero and adjusted the fitting procedure accordingly. Also, certain other lower order terms may be superfluous for the data at hand and can be deleted without decreasing predictive

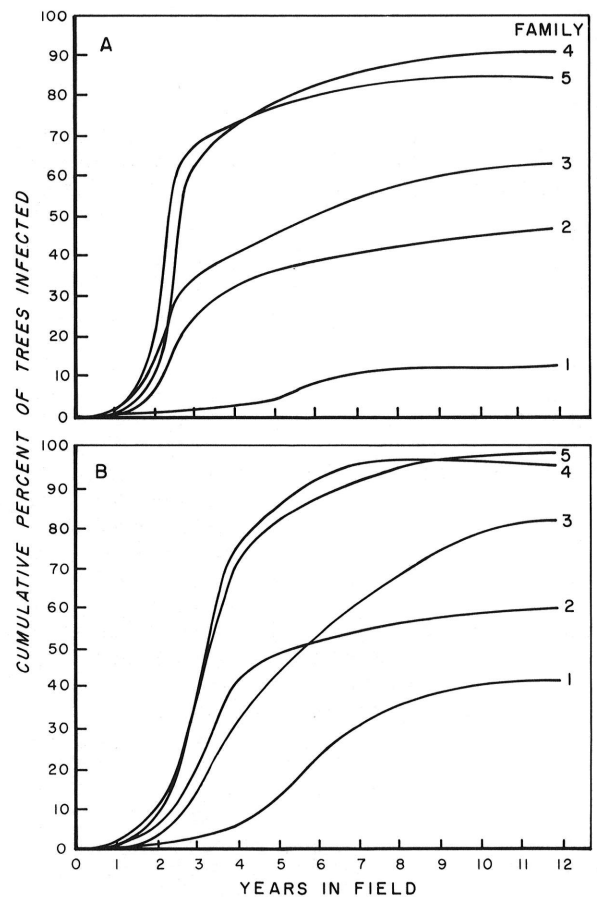


Fig. 1—(A, B). Fusiform rust disease progress curves for open-pollinated progeny of five slash pines at two locations. A) Location 1, and B) Location 2.

capacity. To illustrate, the disease progress curves for each family at Location 1 were fitted to a maximum ninth-degree polynomial (measurements were taken at 10 points in time), while those for each family at Location 2 were fitted to a maximum seventh-degree polynomial (measurements were taken at eight points in time). When the fitted polynomials were compared to actual curves, poor fits often occurred because most curves—especially those for susceptible families—showed dramatic, nonlinear increases in percent infection initially (Fig. 1). Indeed, when polynomials were fitted with zero linear ( $a_1$ ) terms, much lower degree polynomials were required for adequate fits.  $R^2$  values at the fourth degree for each family at each location were 90% or higher with the exception of family 1 at Location 1 where it was 81%. Using polynomials of higher degree gave little or no increase in goodness of fit. Thus, the graphed fourth-degree polynomial curves with  $a_0 = a_1 = 0$ , were in close agreement with the actual curves (Fig. 2). The typical polynomial form appeared as:

$$y(t) = a_2t^2 + a_3t^3 + a_4t^4$$

*Statistical comparison of coefficients.*—Coefficients were used to statistically describe and compare the course of percent infection across families and locations. To simplify later discussion, we let the symbol  $p$  represent the

number of fitted coefficients rather than the degree of the polynomial.

Assume that polynomials of the same form are fitted to each plot. Then there are  $p$  coefficients (=  $a$ 's) for each plot, and these constitute the multivariate data set to be statistically analyzed. Symbolically, the data set may be represented as:

$$p^AN = (a_1 a_2 \dots a_N)$$

where  $N$  (= 60) is the total number of plots,  $p$  is the number of coefficients in each polynomial,  $A$  is a matrix of  $p$  rows and  $N$  columns, and the  $a$ 's are column vectors of length  $p$  containing coefficients of each plot.

A univariate analysis of variance on each coefficient was performed with all effects considered fixed. While the  $p$  univariate analyses provided valuable information on the differences between single coefficients, they did not provide information on the total set of coefficients that defined a curve. We applied multivariate analysis of variance procedures (3, 19) to test a hypothesis concerning differences in entire disease progress curves. The form of the multivariate analysis was the same as above, but  $p \times p$  matrices of mean squares replaced single mean squares, and Wilks' lambda statistics ( $P = 0.05$ ) replaced univariate  $F$ -statistics.

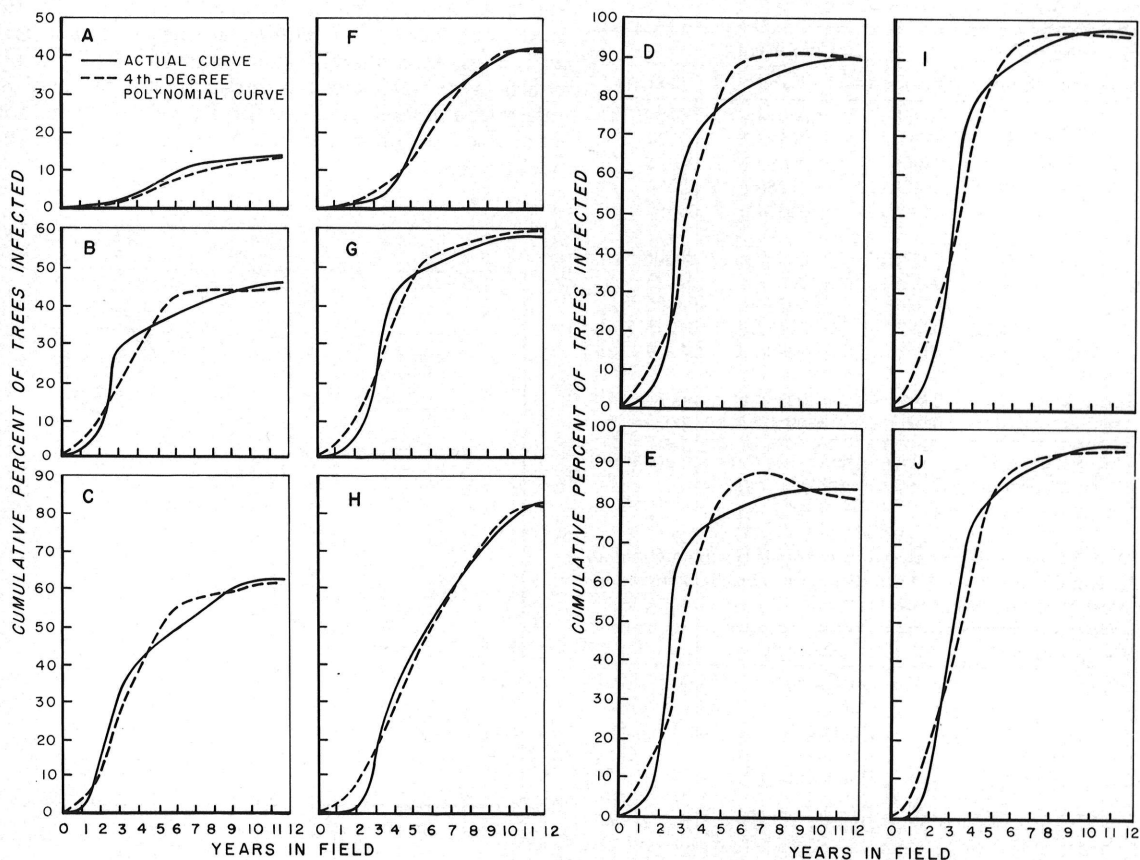


Fig. 2-(A to J). Actual versus 4th-degree polynomial disease progress curves for open-pollinated progenies of five slash pines at Location 1 (A-E) and Location 2 (F-J): A, F) Family 1; B, G) Family 2; C, H) Family 3; D, I) Family 4, and E, J) Family 5.

Wilks' lambda statistics provided a test criterion for testing the hypothesis of no real effects for each of the factors (location, families) and their interactions on the disease progress curves (represented by the polynomial coefficients). If an effect was judged significant, then a discriminant function analysis was applied in order to study the spatial distances between group centroids (multivariate means). These procedures were carried out with the aid of a computer program named MANOVA (2). A detailed discussion of MANOVA may be found in Cramer (3) and Seale (19).

### RESULTS

Generally, the disease progress curves for the five open-pollinated slash pine families resembled those of other plant disease epidemics (22). Three major categories of curve shapes were apparent (Fig. 1 and 2, Table 1): susceptible (families 4 and 5); intermediate (families 2 and 3); and resistant (family 1). The course of infection for susceptible families showed a dramatic increase in percent infection initially which is reflected by large  $a_2$  coefficients. Cumulative infection leveled off 6 to 8 yr

after planting, which is reflected by large negative  $a_3$  and large positive  $a_4$  coefficients. Hence, a relatively short period of time was required for the epidemic to reach high levels. Intermediate families showed a less rapid pattern of disease increase and a lower infection level, reflected by intermediate  $a_2$  coefficient values, smaller negative  $a_3$  and smaller positive  $a_4$  coefficients than those for susceptible families. Disease increase for the resistant family proceeded more slowly and reached even lower levels, as reflected by the very small coefficients compared to susceptible and intermediate families. Thus, genetic resistance evident as early as the second growing season slowed the epidemic.

Regardless of family, disease increased most rapidly between years 2 and 3 at Location 1, and between years 3 and 4 at Location 2. The difference is reflected by the generally larger  $a_2$  coefficients at Location 1 (Table 1). Despite differences in timing, curve shapes did not vary greatly between locations and practically leveled off within 8 to 10 yr after planting at both locations.

Family effects from univariate analysis of variance were by far the dominant factor influencing variation in the three coefficients defining the disease progress curves. Location and location  $\times$  family interaction effects were much weaker, though still significant for two of the three coefficients (Table 2).

These results confirmed that family effects were expressed early; i.e., within the first 3 yr, as evidenced by the large family effect on the quadratic coefficient, the most important one for defining the shape of the curves at early ages. In contrast, location and location  $\times$  family effects were not expressed early. Hence, the level of infection ultimately attained for a family or plantation seemed a function of location or site as well as genetic background.

TABLE 1. Fourth-degree polynomial coefficient terms defining fusiform rust disease progress curves for five slash pine families at two locations when  $a_0 = a_1 = 0$

Location	Family	Coefficients		
		Quadratic	Cubic	Quartic
1	1	.00268 a <sup>z</sup>	-.00014 a <sup>z</sup>	-.0000017 a <sup>z</sup>
	2	.03797 b	-.00578 b	.00024 b
	3	.04606 b	-.00680 b	.00028 b
	4	.08323 c	-.01286 c	.00054 c
	5	.08516 c	-.01324 c	.00055 c
2	1	.00392 a <sup>z</sup>	.00059 a <sup>z</sup>	-.00006 a <sup>z</sup>
	2	.04236 c	-.00598 c	.00023 b
	3	.02646 b	-.00246 b	.00006 a
	4	.07867 d	-.01152 d	.00046 c
	5	.07057 d	-.00996 d	.00039 c
Loc 1 $\bar{x}$	=	.04440 a <sup>z</sup>	-.00776 a <sup>z</sup>	.00032 a <sup>z</sup>
Loc 2 $\bar{x}$	=	.05102 a	-.00587 b	.00022 b

<sup>z</sup>Values followed by the same letter do not differ significantly,  $P = 0.05$ , according to Duncan's multiple range test.

TABLE 2. Univariate analysis of variance for each polynomial coefficient defining the fusiform rust disease progress curves on five slash pine families at two locations

Source of variation	df	Coefficient	F	Prob
Family	4, 40	Quadratic	115.00	<.001
		Cubic	88.13	<.001
		Quartic	68.17	<.001
Location	1, 10	Quadratic	4.55	<.059
		Cubic	5.21	<.046
		Quartic	5.20	<.046
Family $\times$ location	4, 40	Quadratic	2.75	<.041
		Cubic	2.76	<.041
		Quartic	2.46	<.061

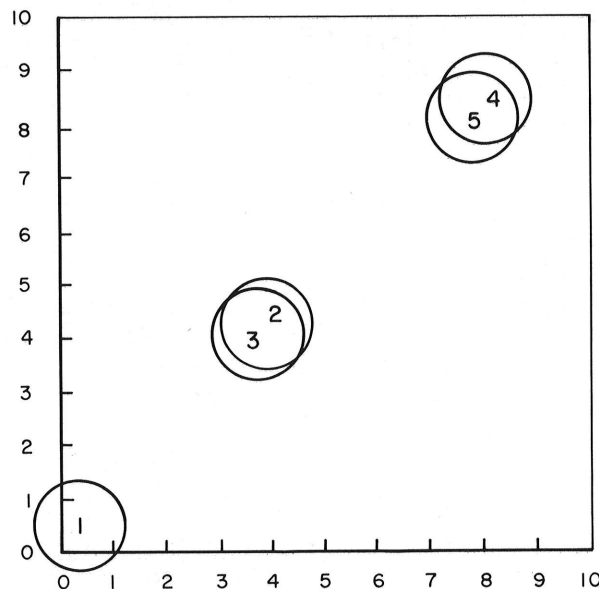


Fig. 3. Plot of first two discriminate function axes for family differences (numbers denote family means and circles denote 0.95 confidence bands about the means).



Duncan's multiple range test of family means overall and by location for all three coefficients largely supported our subjective division of the families into the three categories: "resistant," "intermediate," and "susceptible" (Table 1). Family 3 fitted the intermediate classification at Location 1 but tended to deviate at Location 2. The interaction occurred when less than expected initial infection was followed by much higher than expected later infection at Location 2. This response *may* have been due to growth—family 3 exhibited poor early growth at Location 2; hence, its initial response at Location 2 may have been an indirect result of mediocre growth (4).

Family effects were strong and significant in the joint analysis of all three coefficients by multivariate analysis of variance. However, the location and location  $\times$  family effects were not significant in the joint analysis even though two of the three coefficients were judged significant for these effects in the univariate analysis of variance. The reason for this difference lay in the covariances among the coefficients. They exhibited a very weak pattern of variability for these effects (location and location  $\times$  family). Since multivariate analysis of variance weighted this information for the covariances and variances equally, it was not surprising that the overall effects for location and location  $\times$  family were judged not significant.

A discriminant function analysis of family differences resulted in two significant discriminant function axes which maximally separated family groups. A plot of the resulting spatial separation (Fig. 3) clearly reflected the same categories into which these five families were separated on the basis of univariate analyses in both present and previous reports (4).

#### DISCUSSION

Our analytical approach enabled us statistically to isolate and compare the relative effects of families, locations, and location  $\times$  family interactions on entire fusiform rust disease progress curves. The experiment contained materials with distinct degrees of resistance and, as expected, the genetic (family) effect was the dominant factor influencing the disease progress. Location and location  $\times$  family effects were weaker but, nevertheless, detectable. Results agreed quite closely with those derived from previous univariate analyses of infection at particular points in time. Even though based on differing measurement schedules, infection curves can be compared statistically. This should prove useful to evaluate numerous, large test plantations, or statistically to compare actual and simulated disease progress curves.

The methods described herein may also be useful to analyze other plant disease progress curves because of the method's flexibility. Disease incidence curves under varying edaphic, climatic or biotic conditions could be constructed and then subjected to the analysis presented here. Once the various components are quantified and tested, a more complete model of the disease system could be developed. Such models should prove useful in deriving strategies to manage disease and allocate resources.

#### LITERATURE CITED

1. BEATON, A. E., and J. W. TUKEY. 1974. The fitting of

- power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics* 16(2):147-185.
2. CLYDE, D. J., E. M. CRAMER, and R. J. SHERIN. 1966. Multivariate statistical program. Biometrics Lab., Univ. Miami, Coral Gables, Fla., 61 p.
3. CRAMER, E. M. 1976. A nonorthogonal analysis of variance program. *J. Am. Statist. Assoc.* 71:93-95.
4. DINUS, R. J. 1969. Testing slash pine for rust resistance in artificial and natural conditions. Pages 98-106 in *Proc. 10th South. For. Tree Improv. Conf.*, June 17-19, 1969. Houston, Tex., 235 p.
5. DINUS, R. J., G. A. SNOW, A. G. KAIS, and C. H. WALKINSHAW. 1975. Variability of *Cronartium fusiforme* affects resistance breeding strategies. Pages 193-196 in *Proc. 13th South. For. Tree Improv. Conf.*, June 10-11, 1975, Raleigh, N.C., 262 p.
6. DWINELL, L. D. 1974. Susceptibility of southern oaks to *Cronartium fuiforme* and *Cronartium quercuum*. *Phytopathology* 64:400-403.
7. GRIGGS, M. M., and R. A. SCHMIDT. 1977. Increase and spread of fusiform rust. Pages 32-38 in R. J. Dinus and R. A. Schmidt, eds. *Management of fusiform rust in southern pines*. *Sympos. Proc.*, Univ. Fla., Gainesville, 163 p.
8. HOLLIS, C. A., and R. A. SCHMIDT. 1977. Site factors related to fusiform rust incidence in north Florida slash pine plantations. *For. Sci.* 23:69-77.
9. HOLLIS, C. A., W. H. SMITH, R. A. SCHMIDT, and W. L. PRITCHETT. 1975. Soil and tissue nutrients, soil drainage, fertilization, and tree growth as related to fusiform rust incidence in slash pine. *For. Sci.* 21:141-148.
10. JEWELL, F. F., and S. L. MALLETT. 1967. Testing slash pine for rust resistance. *For. Sci.* 13:413-418.
11. KRANZ, J. 1974. Comparison of epidemics. Pages 355-374 in K. F. Baker, G. A. Zentmeyer, and E. B. Cowling, eds. *Annu. Rev. Phytopathol.* 12. 502 p.
12. KRANZ, J. ed. 1974. *Epidemics of plant diseases; mathematical analysis and modeling (Ecological studies, v. 13)*. Springer-Verlag, Berlin. 170 p.
13. MANION, P. D., and M. BLUME. 1975. Epidemiology of hypoxylon canker of aspen. Abstract NE-42. Page 101 in *Proc. Am. Phytopathol. Soc.* 2. 164 p.
14. MC DONALD, G. I. 1973. Integrated control of white pine blister rust. Abstract 0141 in *Abstracts of papers, 2nd Int. Cong. Plant Pathol.*, Sept. 5-12, 1973. Minneapolis, Minn. (unpaged).
15. MERRILL, W. 1967. The oak wilt epidemics in Pennsylvania and West Virginia: An analysis. *Phytopathology* 57:1206-1210.
16. POWERS, H. R., JR., J. P. MC CLURE, H. A. KNIGHT, and G. F. DUTROW. 1974. Incidence and financial impact of fusiform rust in the South. *J. For.* 72:398-401.
17. RAO, C. R. 1965. The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika* 52:447-458.
18. ROCKWOOD, D. L., and R. E. GODDARD. 1973. Predicted gains for fusiform rust resistance in slash pine. Pages 31-37 in *Proc. 12th South. For. Tree Improv. Conf.*, June 12-13, 1973. Baton Rouge, La. 352 p.
19. SEALE, H. L. 1964. *Multivariate statistical analyses for biologists*. Wiley, New York. 207 p.
20. SNOW, G. A., R. C. FROELICH, and T. W. POPHAM. 1968. Weather conditions determining infection of slash pines by *Cronartium fusiforme*. *Phytopathology* 58:1537-1540.
21. STONECYPHER, R. W., B. J. ZOBEL, and R. L. BLAIR. 1973. Inheritance patterns of loblolly pines from a nonselected natural population. *N.C. Agric. Exp. Stn. Tech. Bull.* 220. 60 p.
22. VAN DER PLANK, J. E. 1963. *Plant diseases: Epidemics and control*. Academic Press, New York. 349 p.

23. WEIDENSAUL, T. C., and F. A. WOOD. 1974. Analysis of a maple canker epidemic in Pennsylvania. *Phytopathology* 64:1024-1027.
24. WISHART, J. 1938. Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika* 30:16-28.