

Sequential Sampling for Estimation and Classification of the Incidence of Hop Powdery Mildew I: Leaf Sampling

David H. Gent, United States Department of Agriculture–Agricultural Research Service (USDA-ARS), Forage Seed and Cereal Research Unit, Oregon State University, Department of Botany and Plant Pathology, Corvallis 97331; **William W. Turechek**, USDA-ARS, United States Horticultural Research Laboratory, Fort Pierce, FL 34945-3030; and **Walter F. Mahaffee**, USDA-ARS, Horticultural Crops Research Laboratory, Oregon State University, Department of Botany and Plant Pathology, Corvallis

ABSTRACT

Gent, D. H., Turechek, W. W., and Mahaffee, W. F. 2007. Sequential sampling for estimation and classification of the incidence of hop powdery mildew I: Leaf sampling. *Plant Dis.* 91:1002-1012.

Hop powdery mildew (caused by *Podosphaera macularis*) is an important disease of hops (*Humulus lupulus*) in the Pacific Northwest. Sequential sampling models for estimation and classification of the incidence of powdery mildew on leaves of hop were developed based on the beta-binomial distribution, using parameter estimates of the binary power law determined in previous studies. Stop lines, models that indicate that enough information has been collected to estimate disease incidence and cease sampling, for sequential estimation were validated by bootstrap simulations of a select group of 18 data sets (out of a total of 198 data sets) from the model-construction data, and through simulated sampling of 104 data sets collected independently (i.e., validation data sets). The achieved coefficient of variation (C) approached prespecified C values as the achieved disease incidence (\hat{p}) increased. Achieving a C of 0.1 was not possible for data sets in which $\hat{p} < 0.10$. The 95% confidence interval of the median difference between the true p and \hat{p} included zero for 16 of 18 data sets evaluated at $C = 0.2$ and all data sets when $C = 0.1$. For sequential classification, Monte-Carlo simulations were used to determine the probability of classifying mean disease incidence as less than a threshold incidence, p_t (operating characteristic [OC]), and average sample number (ASN) curves for 16 combinations of candidate stop lines and error levels (α and β). Four pairs of stop lines were selected for further evaluation based on the results of the Monte-Carlo simulations. Bootstrap simulations of the 18 selected data sets indicated that the OC and ASN curves of the sequential sampling plans for each of the four sets of stop lines were similar to OC and ASN values determined by Monte Carlo simulation. Correct classification of disease incidence as being above or below preselected thresholds was 2.0 to 7.7% higher when stop lines were determined by the beta-binomial approximation than when stop lines were calculated using the binomial distribution. Correct decision rates differed depending on the location where sampling was initiated in the hop yard; however, in all instances were greater than 86% when stop lines were determined using the beta-binomial approximation. The sequential sampling plans evaluated in this study should allow for rapid and accurate estimation and classification of the incidence of hop leaves with powdery mildew, and aid in sampling for pest management decision making.

Hop (*Humulus lupulus*) is a climbing perennial bine grown for its cones, which have been used in the brewing of beer since at least 400 to 600 A.D. (1). The lupulin glands of hop cones produce soft

resins (α and β acids) and essential oils that impart flavor and aroma to beer, and also serve as preservatives. Nearly all commercial hop production in the United States is located in Pacific Northwest states, although small commercial plantings of hop can be found in northern California, the upper Midwest, and the northeastern United States (1,17).

In 1997, hop powdery mildew (caused by *Podosphaera macularis*), one of the oldest and most destructive diseases of hop (26,31), was observed for the first time in commercial hop yards near the Yakima Indian Reservation in Washington (29). Greater than 800 ha of a susceptible cultivar were cut at the ground and abandoned that year because of severe infection of the cones by *P. macularis* (23). The following season, hop powdery mildew was confirmed in all major hop-growing regions in Idaho, Oregon, and Washington, and the

disease now occurs annually in most hop yards planted to susceptible cultivars in these states. Estimated economic impacts of the disease from direct yield and α -acid losses, as well as additional management costs, are estimated at nearly \$740/ha annually (23).

Hop powdery mildew can affect all aerial plant parts, but the most significant economic damage is caused by infection of developing flowers (burs) and cones (26,31). Infected burs and young cones are misshapen or aborted, resulting in direct losses in cone yield and reductions in α -acid content (26,31). Infection of cones close to harvest also is associated with premature senescence that leads to rapid deterioration of cone quality (23,31) or browning during kiln drying (23). Although leaf and stem infections cause little or no direct damage to hop, management of the leaf phase of the disease is essential to protect cones from severe infection because diseased leaves produce the primary inoculum for burr and cone infections later in the season (23,31,37). Young leaf tissue remains susceptible to infection for approximately 2 weeks after emergence (37) and, unlike many hosts (16), hop produces copious amounts of new leaf tissue throughout the season. Consequently, management of the leaf phase of the disease is difficult in hop because fungicides must be applied throughout the season to protect newly emerging leaves and, thus, minimize the risk of cone infection (23,24,37). In practice, scheduling of fungicide applications in a yard is based on disease risk forecasts and the disease status of that yard, as determined by regular scouting (23).

Economic thresholds and economic injury levels for powdery mildew on hop leaves have not been established fully. Nonetheless, a prerequisite for implementation of an action threshold is a sampling procedure to estimate disease incidence or classify disease incidence as above or below some prespecified value. Sequential sampling methods have been used extensively in entomology to minimize sampling costs for assessing pest density (3,12,28,30), but less extensively for assessing foliar diseases of plants (9,10,22,34,35). Depending on the objective, two forms of sequential sampling,

Corresponding author: D. H. Gent
E-mail: gentd@onid.orst.edu

The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the USDA or the ARS of any product or service to the exclusion of others that may be suitable.

Accepted for publication 12 March 2007.

doi:10.1094/PDIS-91-8-1002

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, 2007.

sequential estimation and sequential classification, can be used to assess the disease status of a host population when the spatial pattern of disease is known or can be estimated (4,20). Sequential estimation is used when it is necessary to estimate disease at a prespecified level of precision (20,22). Diseased individuals are tallied from each sampling unit during assessment, and sampling ceases when sufficient information has been collected to estimate disease at the specified level of precision. Often, however, it is not necessary to obtain a precise estimate of disease in a population to make an appropriate pest management decision (3,4,27). In such circumstances, sequential classification can be used to test a hypothesis about disease status, such as whether disease incidence is above or below a critical value important for disease management. The number of sampling units required to classify disease above or below the critical value depends on how much the disease incidence differs from the critical value (20). Disease incidence can be classified with limited sampling when the disease incidence is much higher or lower than the critical value, but a greater number of units must be sampled to classify disease incidence when the incidence is near the critical value. Generally, fewer samples are required for sequential classification than sequential estimation because precision is not controlled for all levels of disease using the former and sampling ceases when enough individuals have been sampled to classify disease incidence.

Development of a rapid and reliable sampling strategy for hop powdery mildew on leaves would enable growers to minimize scouting time and costs, and could be integrated into an overall decision-support system for management of this disease. Therefore, the objectives of this research were to develop and validate sequential sampling plans for (i) estimating the mean incidence of hop powdery mildew on leaves and (ii) classifying the mean incidence of powdery mildew on hop leaves as above or below selected disease threshold values. Part II of this article series addresses similar objectives for sequential sampling for the incidence of powdery mildew on hop cones.

MATERIALS AND METHODS

Field sites and data collection. The incidence of hop powdery mildew on leaves was sampled in commercial hop yards in the primary hop-growing regions of Oregon and Washington from 2000 to 2001 and 2004 to 2005; namely, the Willamette Valley in Oregon and the Yakima Valley and surrounding areas in Washington. All Oregon hop yards included in this study were located in Marion County. Yards in Washington were located in three distinct hop-production regions: the Yakima Indian Reservation, near Moxee City, and the

Lower Yakima Valley (near Mabton and Prosser), as described by Gent et al. (11). The severity of hop powdery mildew varies among these regions, partly because of regional differences in climate and grower cultural practices (1,36), and generally is greatest in and around the Yakima Indian Reservation, followed in order by Moxee City, the Lower Yakima Valley, and Oregon.

Model construction data set. The model construction data set was obtained from an analysis of spatial patterns of hop powdery mildew on leaves (36). Data for model construction was collected from 54 commercial hop yards in 2000 and 2001, in which 198 assessments of disease incidence were conducted. Yards were assessed every 3 to 4 weeks beginning in May and continuing through harvest in late August or September. Estimated mean disease incidence (\hat{p}) was assessed using a cluster sampling design (14). In 2000, yards were stratified into multiple strata (H), where H = number of rows in a yard/20 (rounded up to the nearest integer). A single number, r , between 1 and 20 was chosen randomly and the r th row from each stratum was sampled as described below; this is considered systematic sampling. In 2000, H was variable, and in 2001 H was set at 2. In each row, 10 leaves (n) were selected arbitrarily from first 100 (Oregon) or 75 (Washington) plants (N , sampling units) along the transect, and each leaf was rated for presence or absence of powdery mildew to determine incidence. The total number of leaves assessed from each row was nN , and $n\sum N_H$ leaves were assessed per yard, where N_H is the number of sampling units in the H th stratum. Mean incidence of diseased leaves (\hat{p}) was calculated as $\hat{p} = \sum x/(n\sum N_H)$, where x is the number of diseased leaves per sampling unit. Following the terminology of Turechek and Mahaffee (36), this analysis was conducted at the "yard level" because the sampling units from all rows sampled in a hop yard were used in the calculation of \hat{p} . Therefore, a "yard-level data set" consisted of $n\sum N_H$ leaves evaluated from N sampling units among H strata in a yard during a disease assessment. Yard-level data sets consisted of multiple "row-level data sets," each consisting of nN leaves assessed from one row of a yard. Information on spatial patterns of disease incidence was used to construct the sequential sampling plans described below.

Validation data set. Surveys for hop powdery mildew were conducted in 27 commercial hop yards in Oregon (16 yards) and Washington (11 yards) from 2004 to 2005. Yards were assessed weekly for hop powdery mildew, beginning in May and continuing through August or September. A transect of 100 (Oregon) or 75 (Washington) plants (N) was sampled from each of the first two strata (rows 1 to 40) selected arbitrarily from each yard, as

described above for the model construction data set. Therefore, 1,000 (Oregon) or 750 (Washington) leaves were assessed from each row, and 2,000 (Oregon) or 1,500 (Washington) leaves were assessed from each yard. The first transect always was located closer to the edge (eastern or western side) of the yard than the second transect. From the validation data set, 104 yard-level data sets (208 row-level data sets) were selected randomly to represent two α -acid cultivars (CTZ and Chelan) and two aroma-type cultivars (Glacier and Willamette), the five production regions described above, nine growers, and the range of disease incidence typically found in Oregon and Washington.

Spatial pattern analyses (validation data set). *Distributional analysis.* The beta-binomial and binomial distributions were fit to the incidence data using the computer program BBD (18). For binary data collected as a cluster sample, a good fit to the binomial distribution is an indication of a random pattern of diseased plants, whereas a good fit to the beta-binomial distribution is an indication of an aggregated disease pattern when the parameter θ of the beta-binomial distribution is greater than 0 (13,19). A log-likelihood ratio test statistic was calculated to determine whether the data fit the beta-binomial distribution better than the binomial distribution.

The degree of aggregation of disease incidence was quantified using the parameter θ of the beta-binomial distribution, which provides a measure of variation in disease incidence per sampling unit that generally is referred to as heterogeneity (21,34). Additionally, the index of dispersion, D , was calculated by dividing the observed variance of diseased leaves (v_{obs}) by the theoretical variance for a binomial distribution (v_{bin}) where

$$v_{\text{obs}} = [\sum (x_i - \hat{p}n_i)^2] / (N - 1) \quad (\text{eq. 1})$$

$$v_{\text{bin}} = n\hat{p}(1 - \hat{p}) \quad (\text{eq. 2})$$

and x_i , \hat{p} , n , and N are as defined previously. When $\theta = 0$ or $D = 1$, the pattern of diseased plants is random. Aggregation is indicated when $D > 1$ or $\theta > 0$, with the degree of aggregation directly proportional to the magnitude of the statistic. D has a χ^2 distribution, and can be used to test the null hypothesis of a random distribution of disease incidence with $N - 1$ degrees of freedom (20).

Binary power law analyses. The observed and theoretical binomial variances were fit to the binary form of Taylor's power law (13,33), calculated as

$$\ln(v_{\text{obs}}) = \ln(A_x) + b \ln[n\hat{p}(1 - \hat{p})] \quad (\text{eq. 3})$$

where $\ln(A_x)$ and b are the intercept and slope parameters of a straight line, respectively. When $A_x = 1$ and $b = 1$, equation 1 indicates a random pattern of disease that can be represented by the binomial distribution.

bution. When $A_x > 1$ and $b = 1$, disease incidence is aggregated, but aggregation is not dependent on disease incidence (p). Values of $b > 1$ indicate that aggregation is systematically related to p . Ordinary least squares regression was used to estimate the intercept and slope parameters with SAS (version 9.1, PROC REG; SAS Institute, Cary, NC) after a covariance analysis was conducted to verify that the different sample sizes collected in Oregon ($N = 100$) and Washington ($N = 75$) did not affect the intercept ($p = 0.089$) or slope ($p = 0.878$) of the binary power law. Covariance analysis was conducted in SAS (PROC GENMOD) as described by Gent et al. (11).

Sequential sampling curves. *Sequential estimation.* Statistical methods for development of sequential estimation models for powdery mildew on hop leaves were based on methods presented by Madden and Hughes (20) and Madden et al. (22). Binary power law parameters estimated by regression analysis from the model construction data set were used to develop the sequential estimation models. Precision was expressed in terms of the coefficient of variation, $C = SE \hat{p} / \hat{p}$, where SE is the standard error (21,22). The SE of \hat{p} was expressed in terms of the binary power law parameters as

$$\sqrt{a[\hat{p}(1-p)]^b / N} \quad (\text{eq. 4})$$

where $a = A_x n^{b-2}$ (14,34). This expression for a is based on the binary power law for count data and is different from the expression of a for proportions, where $a = A_p n^{-b}$.

In sequential sampling for estimation, the cumulative number of diseased leaves over N sampling units, T_N , is tallied after each sampling unit is assessed. Sampling ceases when T_N reaches or exceeds a threshold value, referred to as the stop limit, which is defined by a , b , n , N , and C . Disease incidence, as determined by sequential sampling, is then calculated as $\hat{p} = T_N / nN$. The stop lines can be calculated exactly for a binomial distribution (i.e., $\theta = 0$). However, b was >1 for powdery mildew on hop leaves (36) and, under this condition, T_N cannot be written as a simple function of a , b , n , N , and C with a discrete solution (22,34). Equation 4 can be written as a function of $T_N(\gamma_N)$ and solved numerically for T_N as

$$\gamma_N = T_N^{b-2}(nN - T_N)^b = (C^2/a)n^{2b-2}N^{2b-1} \quad (\text{eq. 5})$$

A Mathcad (Mathsoft Inc., Cambridge, MA) worksheet (34) was developed to solve equation 5 iteratively for T_N when $N = 1$ to 500 and $C = 0.1, 0.2$, and 0.3 .

Sequential classification. Statistical methods used for development of the sequential classification models were based on Wald's sequential probability ratio test (SPRT) (39), as modified by Madden and Hughes (20). Reviews of the SPRT have been published (3,4) and only an overview of the methods is presented here. For se-

quential classification, $p_t = (p_0 + p_1)/2$ is defined as some critical value of disease incidence, such as a management threshold for a fungicide application. The variables p_0 and p_1 represent the lower and upper boundaries of disease incidence such that, when the true incidence of disease, p , is equal to or less than p_0 , the field is classified correctly at least $100(1 - \alpha)\%$ of the time. When the true incidence of disease is equal to or greater than p_1 , the field is classified correctly at least $100(1 - \beta)\%$ of the time. In practice, the resulting classifications are interpreted as a test of the null hypothesis $H_0: p \leq p_t$ against the alternative hypothesis $H_1: p > p_t$, respectively. A type I error is made when the disease incidence, p , is incorrectly classified as greater than the critical value, p_t . A type II error is made when p is incorrectly classified as less than p_t . The rate of these two errors is expressed by the operating characteristic (OC), which is defined as the probability of accepting the null hypothesis given the true value of p . If p_t is a disease action threshold, the OC also may be defined as the probability of not making a decision to apply a control measure (4). The OC = $1 -$ (type I error rate) when $p \leq p_t$, and is the type II error rate when $p > p_t$. The OC of a perfect sampling plan is 1 when $p \leq p_t$ and 0 when $p > p_t$. In practice, it is difficult or impossible to achieve such an OC, but the steepness of the OC provides an indication of the error rate of the sampling plan (4). Plots of average sample number (ASN) versus p also are used to evaluate the properties of sequential classification sampling plans (4,34). The ASN is the expected number of sampling units that need to be examined in order to accept or reject the null hypothesis for any true value of p . The OC and ASN are expected values and, therefore, provide average properties of the sampling plan over a large number of sampling bouts, and not necessarily the performance of the sampling plan for a specific hypothesis (20).

Sequential classification requires two lines (stop lines) be calculated to represent p_0 and p_1 . Exact calculation of stop lines for the SPRT is not possible for data described by the beta-binomial distribution, but formulae approximating the stop lines have been developed (20). The general formula for stop lines is

$$i_0 + snN < T_N < i_1 + snN \quad (\text{eq. 6})$$

where sn is the common slope of the stop lines, and i_0 and i_1 are intercepts of the lower and upper stop lines, respectively. Intercept terms are defined by a , b , p_0 , p_1 , and error parameters analogous to type I (α) and type II (β) errors. Formulae for s , i_0 , and i_1 given by Madden and Hughes (20, equations 24a to 24c) and Turechek et al. (34, equations A4 to A6).

Because economic thresholds for hop powdery mildew have not been developed fully, sequential classification plans for

hop powdery mildew on leaves were developed for four levels of p_t : 0.025 ($p_0 = 0.0125$, $p_1 = 0.0375$), 0.05 ($p_0 = 0.025$, $p_1 = 0.075$), 0.10 ($p_0 = 0.05$, $p_1 = 0.15$), and 0.15 ($p_0 = 0.05$, $p_1 = 0.25$). These threshold values were evaluated at each of four combinations of α (0.05 and 0.10) and β (0.05 and 0.10). In total, 16 sampling plans were developed and evaluated by plots of OC and ASN versus p . OC and ASN were calculated by Monte-Carlo simulations using a Fortran program developed by J. P. Nyrop and modified by L. V. Madden (34). Based on results of the simulations, four sequential classification sampling plans were evaluated and validated at each of the four levels of p_t described above, at each of $\alpha = \beta = 0.05$ and $\alpha = \beta = 0.10$, using field data not used in the construction of the model.

Sampling plan evaluation. *Sequential estimation—bootstrapping.* Bootstrapping allows a statistic of interest (e.g., achieved N by sequential sampling for estimation) to be estimated empirically as a frequency distribution from sample data (4). Evaluation of sampling plans by bootstrapping is useful in assessing variability in sampling plan performance among fields or, in this case, over a range of values of disease incidence. Evaluation of the sequential estimation sampling plan was conducted by bootstrap simulation of 18 yard-level data sets, each consisting of two row-level transects, that encompassed the range of \hat{p} observed among the 104 model validation data sets collected from 2004 to 2005. The 104 model validation data sets were classified into six disease incidence categories: $0.01 < \hat{p} < 0.025$, $0.025 \leq \hat{p} < 0.05$, $0.05 \leq \hat{p} < 0.10$, $0.10 \leq \hat{p} < 0.20$, $0.20 \leq \hat{p} < 0.40$, and $\hat{p} \geq 0.40$. Three data sets were selected randomly from each disease incidence class for bootstrap evaluation. Evaluation of data sets where $\hat{p} = 0$ at the yard level were not performed because estimates of C are not possible at this value of \hat{p} and, thus, bootstrap evaluations are not informative for determining performance of the sampling plan. Predicted stop limit curves were determined according to equation 5 with $C = 0.1$ and $C = 0.2$, and pooled yard-level estimates of the binary power law parameters, where $a = 0.198$ and $b = 1.099$ (36). Yard-level estimates of the binary power law parameters were used because these values best characterize the variability of hop powdery mildew on leaves among sampling units within a yard (36).

A software macro was written in Minitab (version 14; Minitab Inc., State College, PA) to conduct the bootstrap analysis as described previously (34). For a given bootstrap simulation, sampling units ($n = 10$ leaves from a plant) were sampled randomly one at a time, with replacement, from among all sampling units from the two row-level data sets collected in a given yard. Sampling ceased when the cumula-

tive number of diseased leaves exceeded the model T_N , or when the number of sampling units equaled the total number of sampling units in the data set (200 in Oregon and 150 in Washington). A minimum of three sampling units was collected before sampling ceased to ensure that a representative sample was collected. The achieved C and moment estimate of p then were calculated. The bootstrap simulation was conducted 100 times for each data set and specified values of C . Nonparametric confidence intervals based on the sign test (22) for the achieved C , the difference between the p of the data set (true p) and achieved p using sequential sampling for estimation, and the achieved N for each of the 18 yard-level data sets were calculated as described previously (22,34).

Sequential estimation—simulated sampling validation. Sampling plans also were validated with independent data by simulated sampling. As indicated by Binns et al. (4), simulated sampling can be used to complement model evaluation conducted by bootstrapping because simulated sampling and bootstrapping provide different insights into model performance. Simulated sequential sampling of the 104 yard-level model validation data sets was performed using a Minitab macro (34). For a given data set, the sampling units were entered into the macro in the same order they were collected in the field. The macro simulated sampling of diseased leaves collected from the sampling units, and tallied the cumulative number of diseased leaves until this value exceeded that of the model T_N calculated from equation 5. If the model T_N was not exceeded during sampling in the first transect, sampling continued into the second transect, which is referred to as the “original order.” Simulated sampling of each yard also was conducted by initiating sampling at the second transect, referred to as the “reversed order.” Therefore, each yard was sampled two times (in the original and reversed orders) to determine if the position in the yard at which sampling was initiated affected results of the sequential sample. Yard-level estimates of p , the achieved C , and the achieved N were calculated for both the original and reversed orders.

Sequential classification—bootstrapping. The four sequential classification sampling plans selected after Monte-Carlo simulation were evaluated by bootstrap simulation, as described above for sequential estimation. Stop lines were determined using equation 6 with s , i_0 , and i_1 calculated as reported by Madden and Hughes (20, equations 24a to 24c) and Turechek et al. (34, equations A4 to A6). A Minitab macro was written to conduct the bootstrap analysis as described previously (34). For a given bootstrap simulation, sampling units ($n = 10$ leaves from a plant) were sampled randomly one at a time, with replacement, from among all sampling units from the

yard. Sampling ceased when the cumulative number of diseased leaves exceeded the upper or lower stop lines of the model, or the data set was sampled fully. A minimum of three sampling units was collected before sampling ceased to ensure a representative sample was collected. The achieved OC and ASN then were calculated. The bootstrap simulation was conducted 100 times for each data set and specified values of p_i , p_0 , p_1 , α , and β .

Sequential classification—simulated sampling validation. The four sequential classification sampling plans selected after Monte-Carlo simulation were evaluated by simulated sampling, as described above for sequential estimation. Simulated sampling of 104 yard-level model validation data sets was performed using a Minitab macro (34). For each of the 104 yard-level data sets, the sampling units were entered into the macro in the original order and then in the reversed order of sampling, as described above. Sampling ceased in the first transect if the cumulative number of diseased leaves exceeded the upper or lower stop lines of the model and a decision could be made. If a decision was not made before the end of the first transect, sampling continued into the second transect until a decision could be made or the data set was fully sampled (i.e., 200 sampling units for yards in Oregon and 150 for yards in Washington). Additionally, the 208 row-level data sets were sampled using the macro to validate the sampling plan for individual transects in the manner described for yard-level data sets.

Stop lines were calculated based on either the binomial distribution (34, equations A1 to A3) or beta-binomial distribution (34, equations A4 to A6). To determine whether a correct decision for a data set was made, it was assumed that p for a row-level or yard-level data set represented the true p for that yard. The true value of p for a data set was compared with the hypothesized p_i to determine whether to reject the H_0 in favor of the H_1 . The decision based on results of the simulated sequential classification was compared with the correct decision for the yard, in order to calculate type I and type II error rates. When sampling row-level data sets, a correct decision was recorded if the true p of the row was correctly classified as above or below p_i . At the yard level, where multiple row-level data sets were assessed from each yard, a correct decision at the yard-level was recorded if mean disease incidence was correctly classified above or below p_i for both transects at the row level. A type I error was recorded if mean disease incidence was incorrectly classified as greater than p_i for at least one of the two data sets at the row level, and a type II error was recorded if mean disease incidence was incorrectly classified as less than p_i for at least one of the two data sets at the row level. Type I

and II errors were recorded when mean disease incidence was incorrectly classified as greater than p_i in one row-level data set and less than p_i for the other row-level data set in that yard.

RESULTS

Spatial pattern analyses. A detailed analysis of the model construction data set was published previously (36), and a summary of the results is presented here. Disease incidence for the model construction data sets ranged from 0.0004 to 0.71, with a median of 0.012. For any single data set, the values obtained were assumed to be the true p for that yard. A log-likelihood ratio test indicated that the beta-binomial distribution provided a better fit than the binomial distribution for 47% of the 198 data sets at the yard level. The parameter θ of the beta-binomial distribution ranged from 0 to 0.31, with a median of 0.022, which indicated a low degree of small-scale aggregation of disease. Covariance analysis indicated that the year in which samples were collected did not have a significant effect on the slope (b) and intercept ($\ln[A_x]$) parameters of the binary power law. Therefore, the data were pooled over the years 2000 to 2001, resulting in parameters $\ln(A_x) = 0.457$ (SE = 0.023) and $b = 1.099$ (SE = 0.008). With these parameters, $A_x = 1.579$ and $a = 0.198$.

For the 104 yard-level validation data sets collected in 2004 and 2005, \hat{p} ranged from 0.0005 to 0.65, with a median of 0.022 (Fig. 1A). The parameter θ of the beta-binomial distribution ranged from 0 to 0.71, with a median of 0.0361 (Fig. 1B). The beta-binomial distribution fit better than the binomial distribution for 54.8% of the data sets according to the log-likelihood ratio test. The binary power law provided an excellent fit to the data ($R^2 = 0.981$) (Fig. 1C). Parameter estimates were $\ln(A_x) = 0.621$ (SE = 0.0409) and $b = 1.136$ (SE = 0.0152). With these parameters, $A_x = 1.861$, $a = 0.255$, and $A_p = 3.481$. Heterogeneity in the model construction and validation data sets were similar numerically, although the slope and intercept parameter estimates of the binary power law were significantly greater in the model validation data set than the model construction data set, as determined by t tests (slope $p < 0.0001$ and intercept $p = 0.0083$).

Sequential estimation. Sequential estimation stop limits for estimating mean incidence of powdery mildew on leaves and the corresponding incidence of disease are shown in Figure 2A and B, respectively, for $C = 0.1$, 0.2, and 0.3. The achieved level of precision increased as the true level of disease incidence decreased, or the number of diseased leaves that needed to be sampled in order to end the sequential sampling increased.

Bootstrapping validation. The achieved C was greater than the prespecified C for most data sets where $\hat{p} \leq 0.025$ (Fig. 3A

and B). The achieved C approached the prespecified C as \hat{p} increased, and the median achieved C was less than the prespecified C for 8 of 15 (53%) data sets where $\hat{p} > 0.025$ and $C = 0.2$. The achieved C was less than the prespecified C for 50% of data sets where $\hat{p} > 0.2$ (data sets 13 to 18; Fig. 3A and B) because the data sets tended to be oversampled due

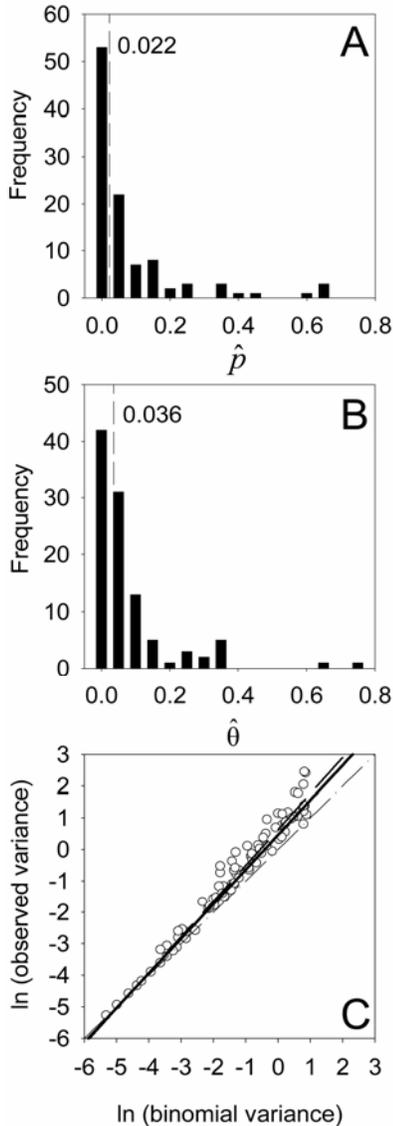


Fig. 1. Frequency distribution of the beta-binomial distribution parameters **A**, \hat{p} and **B**, $\hat{\theta}$ and **C**, the relationship between the natural logarithms of the theoretical variance for a binomial (random) distribution and the observed variance for the incidence of hop powdery mildew on leaves for the model validation data set. **A** and **B**, The vertical dashed lines are the median values for the indicated statistic, with the numerical value shown on the graph. **C**, The dash-dot line represents the theoretical binomial line, the solid line represents the least squares regression fit to the yard-level model construction data set described by Turechek and Mahaffee (36), and the dashed line represents the least squares regression fit to the model validation data set collected from hop yards in Oregon and Washington in 2004 and 2005. The open circles are individual values for 104 yard-level model validation data sets.

to the minimum sampling rule. When C was prespecified at 0.1, the achieved C approached 0.1 for $\hat{p} > 0.10$ (Fig. 3B). For $\hat{p} < 0.05$ (data sets 1 to 6), $C = 0.1$ could not be attained even when the number of plants sampled equaled the total number of sampling units in the data sets (200 for Oregon and 150 for Washington; Fig. 3B and F).

The 95% confidence interval for the median of the difference between \hat{p} and p included zero for all 18 data sets at $C = 0.1$ and 16 of 18 (88.9%) data sets at $C = 0.2$ (Fig. 3C and D), indicating that \hat{p} was close to p for these data sets. Confidence intervals generally increased with increasing p because the sample size (N) decreased with increasing p for both levels of C (Fig. 3E and F). Sample size was greatest for data sets where $p \leq 0.025$ (data sets 1 to 3), and in these data sets the estimated $\hat{p} \approx p$.

Simulated sampling validation. Median values of the difference between p and \hat{p} were 0 for $C = 0.2$ and $C = 0.1$ when the transects were sampled in the original or reversed orders (Fig. 4A and B). The inter-

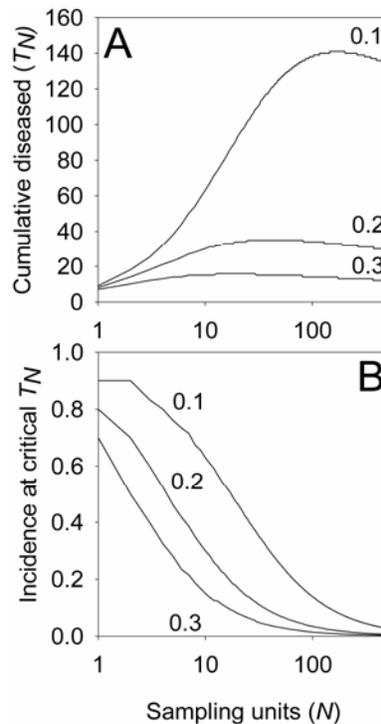


Fig. 2. Sequential-estimation stop limits for estimating the mean incidence of hop powdery mildew on leaves with a coefficient of variation (C) of 0.1, 0.2, and 0.3 indicated numerically on the graph. **A**, Cumulative number of diseased leaves (T_N) versus the total number of sampling units (N) based on equation 2 in the text, with binary power law parameters of $b = 1.099$ and $a = 0.198$. **B**, Mean disease incidence at critical T_N (the point where the observed cumulative number of diseased leaves crosses the model T_N curve) in relation to N . Sampling ceased when the cumulative number of diseased leaves crossed the critical T_N in **A**, at which point the mean disease incidence was then calculated as T_N/nN .

quartile range of $p - \hat{p}$ was 0.011 and 0.004 for $C = 0.2$ during sampling of data sets in the original and reversed orders, respectively, and 0 for $C = 0.1$ for sampling of data sets in the original or reversed orders. As noted by Turechek et al. (34), the difference between p and \hat{p} may be misleading because, when disease incidence is low (< 0.1), the data sets were fully sampled and, thus, $p = \hat{p}$. Over the 104 validation data sets, the difference between p and \hat{p} was > 0.10 in only 10 or 8 data sets when $C = 0.2$ and sampling was conducted in the original or reversed orders, respectively; and 6 or 4 data sets when $C = 0.1$ and transects were sampled in the original or reversed orders, respectively.

Median values of the achieved C were 0.259 and 0.210 for sampling of transects in the original order, which was greater than the prespecified values of $C = 0.2$ and $C = 0.1$, respectively (Fig. 4C and D). Medians values of the achieved C were similar for sampling in reversed order, with values of 0.248 and 0.220 at $C = 0.2$ and $C = 0.1$, respectively. The distribution of the achieved C for both prespecified levels of C was positively skewed in the original order, with interquartile ranges of 0.257 ($Q_1 = 0.194$ and $Q_3 = 0.448$) and 0.210 ($Q_1 = 0.135$ and $Q_3 = 0.448$) for $C = 0.2$ and 0.1, respectively (Fig. 4C and D). The median, first quartile, and third quartile for sampling in reversed order were 0.248, 0.186, and 0.448 at $C = 0.2$, respectively, and 0.220, 0.130, and 0.448 at $C = 0.1$, respectively. Among the 104 model validation data sets, the achieved C was equal to or less than the prespecified C for 30 (original) and 36 (reversed) data sets at $C = 0.2$, and 10 (original) and 14 (reversed) data sets at $C = 0.1$, respectively. For 39 data sets at $C = 0.2$ and 78 data sets at $C = 0.1$, the prespecified value of C could not be attained because all sampling units in the data set were collected and sampling was forced to cease. For those data sets for which the prespecified C was attained, the number of sampling units (N) collected decreased as the p of the yard increased, independent of the order of sampling of the transects (Fig. 4E and F). The minimum number of sample units, imposed by the sampling rule, were collected in four (original) or five (reversed order) data sets at $C = 0.2$, and one data set (original or reversed) at $C = 0.1$. The number of sampling units required to achieve the desired C was nearly identical to the theoretical N in fixed sampling (22, equation 17b) when p was greater than approximately 0.10 and the data set was not sampled entirely.

Sequential classification. Sequential classification stop lines for four combinations of α and β at $p_0 = 0.0125$ and $p_1 = 0.0375$ ($p_t = 0.025$) are shown in Figure 5A. Varying α and β had little effect on the distance between the stop lines. The OC

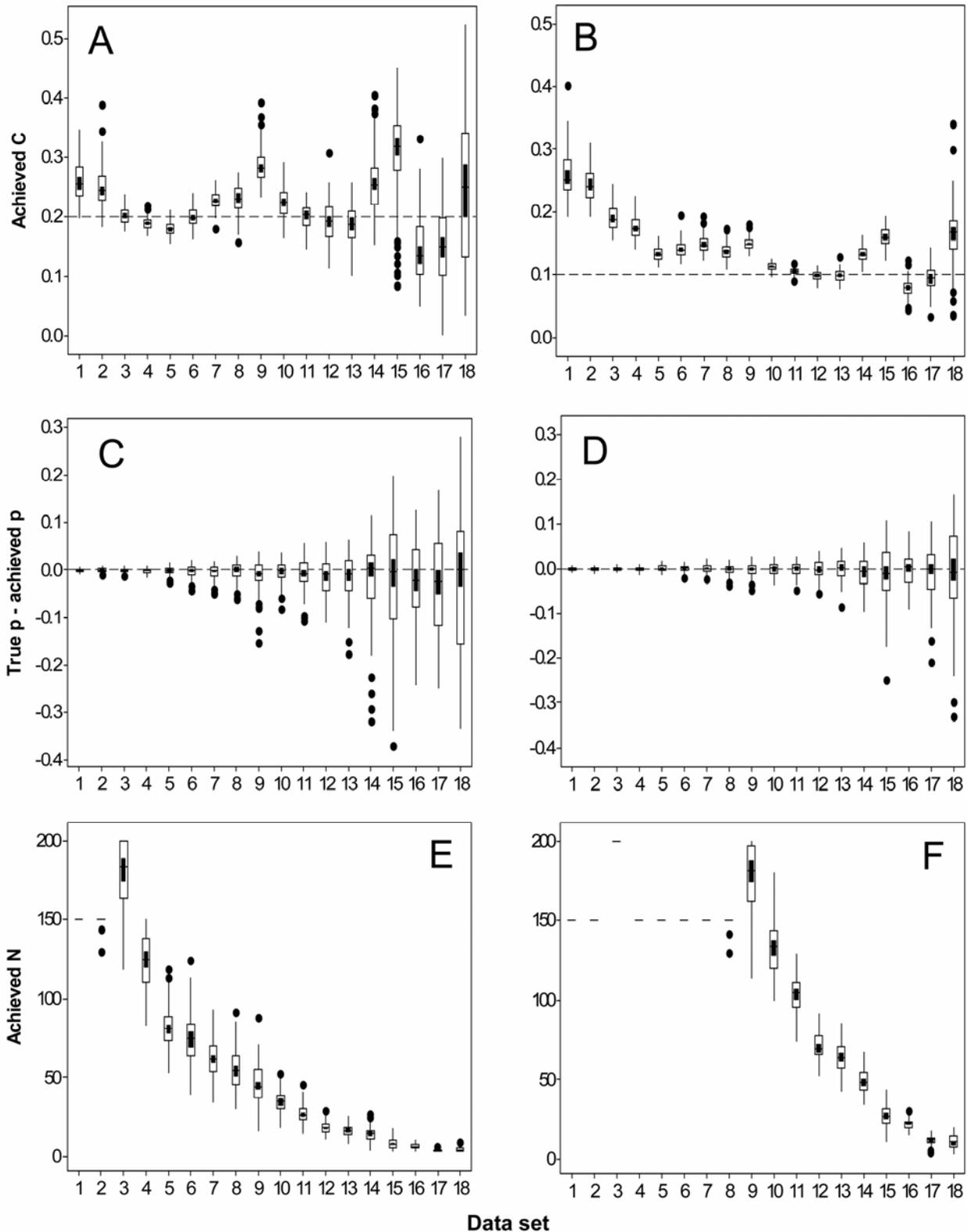


Fig. 3. **A** and **B**, Box plots of the achieved coefficient of variation (C); **C** and **D**, the difference between the true incidence of hop leaves with powdery mildew, p , and the estimated incidence of powdery mildew, \hat{p} , based on sequential samples; and **E** and **F**, the achieved sample size (N) for estimating the incidence of hop powdery mildew on leaves from 100 bootstrap samplings of 18 data sets for model evaluation. Preselected values of C are **A**, **C**, and **E**, 0.2 and **B**, **D**, and **F**, 0.1, and are indicated by dashed lines in **A** and **B**. Sequential estimation stop lines were generated according to equation 2 in the main text, with binary power law parameters of $a = 0.198$ and $b = 1.099$ (36). The 18 model validation data sets were chosen by selecting three data sets randomly from each of seven disease incidence classes ($0.01 < \hat{p} < 0.025$, $0.025 \leq \hat{p} < 0.05$, $0.05 \leq \hat{p} < 0.10$, $0.10 \leq \hat{p} < 0.20$, $0.20 \leq \hat{p} < 0.40$, and $\hat{p} > 0.40$) from among the 198 model construction data sets collected in hop yards in 2000 and 2001. The 18 model evaluation data sets are arranged in ascending order of \hat{p} . Data sets collected in Washington contained 150 sampling units (75 per transect), and data sets collected in Oregon contained 200 sampling units (100 per transect). Box plots show the median (line in open boxes), middle 50% of the data (open box), 95% confidence interval for the median based on the non-parametric sign-test (solid bar inside box) (32), extremes of the data points (whiskers), and outliers (solid circles).

curves were similar among the four combinations of α and β , although the OC was slightly steeper for $\beta = 0.05$ compared with $\beta = 0.10$ when $p > p_t$ (Fig. 5B). That is, a higher correct decision rate was achieved when $p > p_t$. The OC was 0 for all combinations of α and β at $p \geq 0.10$. Varying α and β did have a large

effect on the ASN curves, which was particularly evident when p was near p_t . The ASN (rounded up to the nearest integer) at $p_t = 0.025$ was 51, 45, 42, and 32 sampling units for $\alpha = \beta = 0.05$, $\alpha = 0.10$ and $\beta = 0.05$, $\alpha = 0.05$ and $\beta = 0.10$, and $\alpha = \beta = 0.1$, respectively (Fig. 5C). The ASN was within 1 sampling unit, or iden-

tical, for all combinations of α and β when p was far (≥ 0.10) from p_t .

Changing p_t had a greater effect on the OC and ASN curves than did altering α or β (Fig. 6). Increasing p_t increased the slope of the stop lines (Fig. 6A), with a slight shift of the OC curve to the right (Fig. 6B).

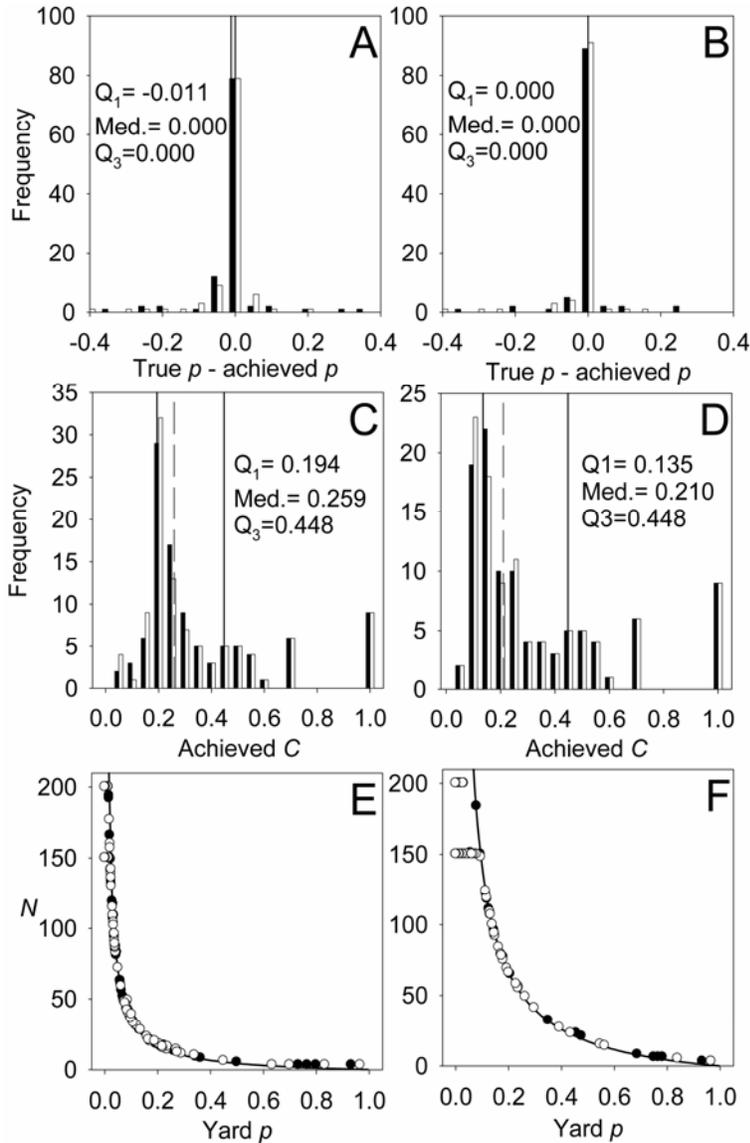


Fig. 4. Results of sequential sampling for validating estimates of the incidence of powdery mildew on hop leaves from 104 yard-level data sets collected from hop yards in Oregon and Washington in 2004 and 2005. **A** and **B**, Frequency distribution of the difference between true disease incidence (p) of the data set and estimated disease incidence from the sequential sample, \hat{p} . **C** and **D**, Frequency distribution of the achieved coefficient of variation. **E** and **F**, Relationship between mean disease incidence for the yard (p) and number of sampling units (N) collected. Preselected values of C are **A**, **C**, and **E**, 0.2 and **B**, **D**, and **F**, 0.1. The solid bars are results from simulated sampling of two transects per yard in the same order in which data was originally collected in the yard (original order). The open bars are results when sampling order of the transects was reversed (reversed order). The dashed lines in **A**–**D** are the median of 104 data sets when sampling was conducted in the original order, and the solid lines are the first and third quartiles for the specified statistic (indistinguishable in **A** and **B**). Numerical values of the statistics are given on the graph. When the row-level data sets were sampled in reverse order, values of the first quartile, median, and third quartile for $p - \hat{p}$ were **A**, -0.004 , 0, and 0, respectively at $C = 0.2$ and **B**, 0, 0, and 0, respectively at $C = 0.1$. Values of the first quartile, median, and third quartile for the achieved C were **C**, 0.186, 0.248, and 0.448, respectively, at $C = 0.2$ and **D**, 0.130, 0.220, and 0.448, respectively at $C = 0.1$. The data sets for which the achieved at $C = 1$ and 0.705 represented nine and six yards, respectively, where 1 diseased leaf was observed from among $n = 2,000$ (Oregon) or $n = 1,500$ leaves (Washington) sampled. **E** and **F**, Solid lines represent the expected number of sampling units given p as calculated by Madden et al. (22, equation 17b). Solid circles are data sets sampled in the original order, and open circles are data sets sampled in reversed order.

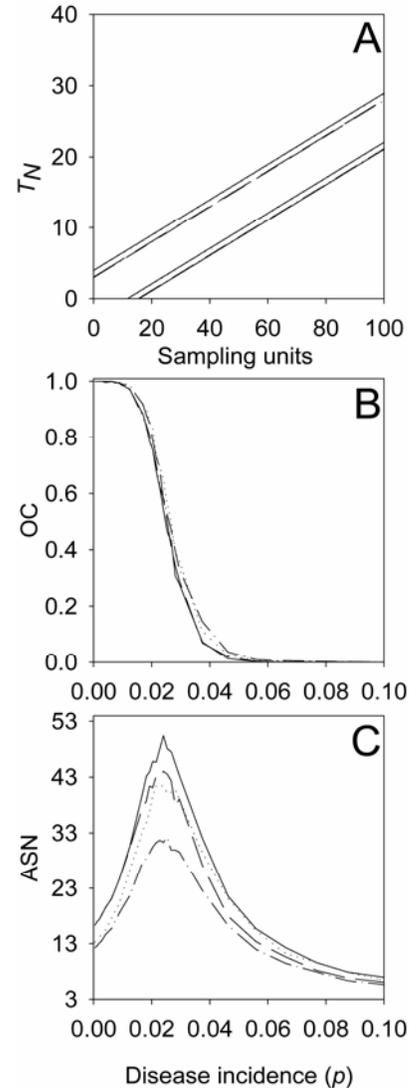


Fig. 5. Stop lines for classifying the incidence of powdery mildew on hop leaves as above or below a critical threshold, p_t , from a cluster sample with $n = 10$ leaves per sampling unit in hop yards sampled in Oregon and Washington in 2004 and 2005. **A**, Stop lines based on Wald's sequence probability ratio test (39) with $p_0 = 0.0125$ and $p_1 = 0.0375$ ($p_t = 0.025$) for four combinations of α and β , where α and β are analogous to type I and type II error rates, respectively. **B**, Operating characteristic (OC) curves, and **C**, average sample number (ASN) curves for sequential classification stop lines defined in **A**. Error probabilities are set at $\alpha = 0.05$ and $\beta = 0.05$ (solid line), $\alpha = 0.10$ and $\beta = 0.05$ (long dashed line), $\alpha = 0.05$ and $\beta = 0.10$ (dotted line), and $\alpha = 0.10$ and $\beta = 0.10$ (dash-dot line). OC and ASN curves were determined by 1,000 Monte Carlo simulations using a modification of the algorithm described by Hoffman et al. (12), with the parameter θ as a function of disease incidence according to the binary power law (described in the main text and by Turecheck et al. [36]), where $a = 0.198$ and $b = 1.099$.

Increasing p_t also flattened the OC curve, indicating an overall increase in the incorrect decision rate (Fig. 6B). Flattening of the OC curve also occurred when the difference between p_0 and p_1 increased. The ASN required to classify disease incidence above or below p_t was reduced by increasing p_t (Fig. 6C). At $p = p_t = 0.025, 0.05, 0.10,$ or 0.15 , ASN (rounded up to the nearest integer) was 51, 33, 21, or 9, respectively.

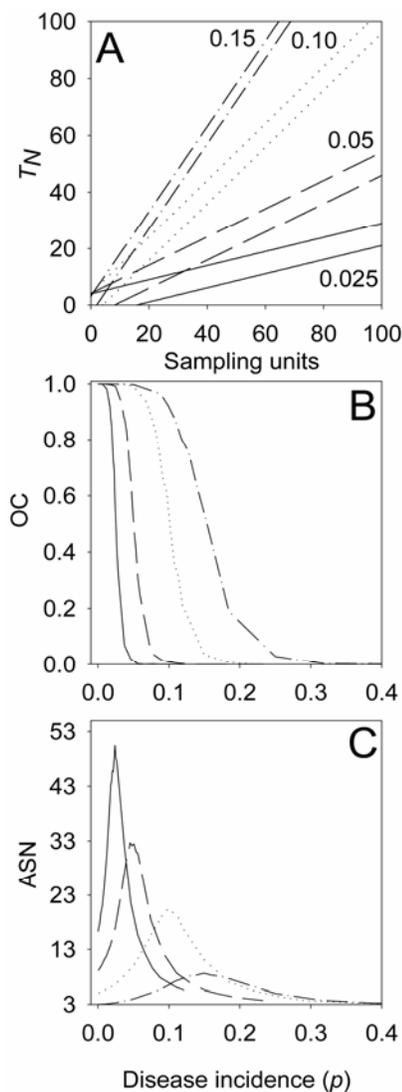


Fig. 6. **A**, Stop lines, **B**, operating characteristic (OC) curves, and **C**, average sample number (ASN) curves for sequential classification as determined by Wald's sequential probability ratio test (39) for the incidence of powdery mildew on hop leaves from a cluster sample with $n = 10$ leaves per sampling unit, in hop yards sampled in Oregon and Washington in 2004 and 2005. OC and ASN curves were determined by 1,000 Monte Carlo simulations using a modification of the algorithm described by Hoffman et al. (12), with the parameter θ described as a function of disease incidence according to the binary power law (described in the main text and by Turecheck et al. [36]), where $a = 0.198$ and $b = 1.099$. Error probabilities in all cases were set at $\alpha = \beta = 0.05$. The critical value, p_t , is written adjacent to the corresponding stop lines in A. The parameters p_0 and p_1 of the stop lines are given in the text.

Bootstrapping validation. Four sequential classification sampling plans were selected for further evaluation by bootstrap simulation based on the results of the Monte Carlo analysis. The parameters of these sampling plans were $p_t = 0.025$ and $p_t = 0.10$ at error rates of $\alpha = \beta = 0.05$ and $\alpha = \beta = 0.10$ (Fig. 7A to D). The achieved OC was very similar to the OC curves obtained by Monte Carlo simulation for the 18 data sets at all combinations of p_t and error rates (Fig. 7A and B). In one data set where $p = 0.079$ ($p = 0.117$ and 0.041 at the row levels), the achieved OC (0.590) was lower than the OC predicted by Monte Carlo simulation (0.785) for error rates of $\alpha = \beta = 0.10$, and the achieved ASN (7.15) also was lower than the simulated ASN (11.07). The difference between the achieved and predicted OC and ASN values occurred because the θ (at the yard level) was 0.373 and much greater than the θ predicted by the binary power law (0.062) (36). Yard-level heterogeneity in this data set was the greatest among all model construction data sets (36, Fig. 1D).

Similarly, for $p_t = 0.10$ and $\alpha = \beta = 0.10$, the ASN for one yard ($p = 0.107$) was substantially lower (13.78) than the ASN derived by simulation (19.47). For this yard, $\theta = 0.129$ and the predicted $\theta = 0.0675$.

Simulated sampling validation. Correct decision rates were 2.0 to 3.9% higher at the row level when stop lines were developed based on the beta-binomial distribution rather than the binomial distribution (Table 1). For the beta-binomial approximation, correct decisions were made in 93.8 and 96.2% of the data sets where $p_t = 0.025$ for $\alpha = \beta = 0.10$ or $\alpha = \beta = 0.05$, respectively. For $p_t = 0.10$, correct decisions were made in 95.2 or 95.7% of the data sets for $\alpha = \beta = 0.10$ or $\alpha = \beta = 0.05$, respectively. For stop lines derived from beta-binomial approximation, type II errors were more common than type I errors with any combination of α , β , and p_t . The higher rate of correct decisions with the beta-binomial-based stop lines resulted in a greater mean and median number of samples being collected compared with the binomial-based stop lines. Depending on

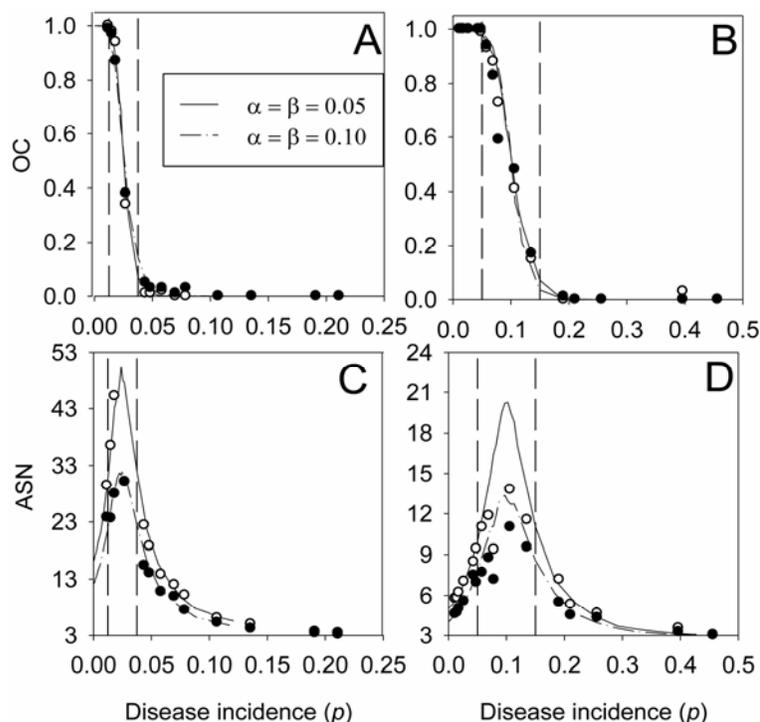


Fig. 7. **A** and **B**, Operating characteristic (OC), and **C** and **D**, average sample number (ASN) curves for sequential classification sampling plans as determined by Wald's sequential probability ratio test (39) for the incidence of powdery mildew on hop leaves from data sets collected at the yard-level in hop yards of Oregon and Washington in 2004 and 2005. Threshold values are **A** and **C**, $p_0 = 0.0125$, $p_1 = 0.0375$, and $p_t = 0.025$; and **B** and **D**, $p_0 = 0.05$, $p_1 = 0.15$, and $p_t = 0.10$. OC and ASN curves were determined by 1,000 Monte Carlo simulations using a modification of the algorithm described by Hoffman et al. (12), with the parameter θ a function of the mean disease incidence according to the binary power law (described in the main text and by Turecheck et al. [36]), where $a = 0.198$, and $b = 1.099$. Curves with error probabilities of $\alpha = \beta = 0.05$ (solid line) and $\alpha = \beta = 0.10$ (dashed line) are shown. Circles are the achieved OC and ASN from 100 bootstrap simulations of sequential sampling for classification of 18 yard-level model validation data sets; open circles are simulations where $\alpha = \beta = 0.05$, and solid circles are simulations where $\alpha = \beta = 0.10$. The 18 model validation data sets were chosen by selecting three data sets randomly from each of six disease incidence classes ($0.010 < \hat{p} < 0.025$, $0.025 < \hat{p} < 0.050$, $0.050 < \hat{p} < 0.100$, $0.100 < \hat{p} < 0.200$, $0.200 < \hat{p} < 0.400$, and $\hat{p} > 0.400$) from among 104 model validation data sets collected during 2004 and 2005. Data is not presented for $\hat{p} > 0.25$ when $p_t = 0.025$, or $\hat{p} > 0.50$ when $p_t = 0.10$, because the OC = 0 and ASN \approx 3 in these cases.

α , β , and p_t , two to five more samples were collected on average before p was classified as above or below p_t . As expected, mean N increased when error rates were lowered. Mean N was 1.9 or 1.3 more when error rates were controlled at 0.05 compared with 0.10 (Table 1).

Correct decision rates were lower when the row-level data sets from a yard were considered as two individual data sets, and a correct decision in both row-level data sets was necessary to make a correct decision at the yard level (Table 1). For the beta-binomial approximation for $p_t = 0.025$, correct decisions were made in 87.5 and 92.3% of data sets at the yard level compared with 93.8 and 96.2% at the row level for $\alpha = \beta = 0.10$ and $\alpha = \beta = 0.05$, respectively. For $p_t = 0.10$, correct decisions were made in 91.3 and 92.3% of data sets when $\alpha = \beta = 0.10$ and $\alpha = \beta = 0.05$, respectively. Type I and type II error rates were higher at the yard level than the row level because correct decisions in both row-level data sets were needed for a correct decision, but were lower than the prespecified error rates for $p_t = 0.025$ and 0.10. At $p_t = 0.10$ and $\alpha = \beta = 0.10$ in one data set (1% of all data sets), a type I error was made in one row-level data set and a type II error was made in the other row-level data set from that yard. In this yard, p in the row-level data sets was near p_t (0.18 and 0.089). A type II error occurred in row 1 when the $p = 0.18$ was classified as less than p_t and a type I error occurred in row 2 when $p = 0.089$ was classified as greater than p_t .

The order in which transects were sampled affected error rates. For $p_t = 0.025$ and error rates controlled at 0.05, the correct decision rate was 7.7% greater when sampling was initiated in the original order in which the data was collected in the yards, compared with when the order of sampling was reversed (Table 2). Conversely, for $p_t = 0.10$, correct decision rates for the yard were 4.2 to 6.7% greater when sampling was initiated in the reversed order compared with the original order. Among the 104 model validation data sets, the mean p in the first and second transects was 0.0788 (median 0.0183) and 0.0833 (median 0.0177), respectively, indicating that the disease incidence, on average, was slightly greater in the interior of yards than closer to the edges of the yards. This may explain why the correct decision rate was greater for $p_t = 0.025$ in the original order of sampling and greater for $p_t = 0.10$ for the reversed order of sampling.

Where higher correct decision rates were observed during sampling data sets in the original versus reversed orders, the increase in correct decision rates generally was associated with a reduction in the rate of type II errors, although the mean and median N were similar (less than 1 sampling unit difference) or identical for sampling transects in the original or reversed orders. Type I errors were controlled at the prespecified rates for all sampling plans; however, prespecified rates of the type II errors were achieved only for $p_t = 0.10$ and $\alpha = \beta = 0.10$ when

transects were sampled in the reversed order (Tables 1 and 2).

DISCUSSION

Management of the foliar phase of hop powdery mildew is critical for successful management of the disease on cones and to minimize crop loss (31,37). Economically sustainable management of the disease requires growers to make management decisions by integrating knowledge of host susceptibility, environmental conditions favorable for the disease, and inoculum density (23,24,36,37). The sequential sampling plans developed in this research, based on knowledge of spatial heterogeneity of the disease, provide the foundation for efficient sampling of hop powdery mildew on leaves and linkage of sampling protocols to thresholds for disease management.

Evaluation of sequential estimation sampling plans by bootstrap analysis of selected data sets from the model construction data, and simulated sampling of data collected independently of the model construction data, revealed that it is possible to estimate disease incidence with reasonable precision ($C = 0.2$) for yards in which the incidence powdery mildew is $p \geq 0.025$. Deviation of the median value of the difference in the true incidence of symptomatic leaves, p , and the incidence of symptomatic leaves estimated by sequential sampling, \hat{p} , tended to increase with p , suggesting some bias in the sequential estimation, as reported previously (4,22,34).

Table 1. Correct decision and error rates (%) of sequential sampling plans for classifying the incidence of powdery mildew on hop leaves for 208 row-level and 104 yard-level data sets collected from hop yards in Oregon and Washington in 2004 and 2005^a

Parameters ^e	Row level ^b				Yard level ^b			
	$p_t = 0.025^c$		$p_t = 0.10^d$		$p_t = 0.025^c$		$p_t = 0.10^d$	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Binomial								
Correct decision	93.3	91.8	91.8	91.8	86.5	83.7	84.6	84.6
Type I error	2.9	2.9	4.3	4.3	5.8	5.8	7.7	7.7
Type II error	3.8	5.3	3.8	3.8	7.7	10.6	6.7	6.7
Type I and II error	0	0	1.0	1.0
Mean N	12.5	9.4	3.8	3.7	12.5	9.4	3.8	3.7
Median N	12.0	9.0	3.0	3.0	12.0	9.0	3.0	3.0
Beta-Binomial								
Correct decision	96.2	93.8	95.7	95.2	92.3	87.5	92.3	91.3
Type I error	1.4	1.9	1.4	1.4	2.9	3.8	1.9	1.9
Type II error	2.4	4.3	2.9	3.4	4.8	8.7	4.8	5.8
Type I and II error	0	0	1.0	1.0
Mean N	17.5	12.5	7.1	5.7	17.5	12.5	7.1	5.7
Median N	16.0	12.0	5.0	4.0	16.0	12.0	5.0	4.0

^a Type I error indicates that mean disease incidence was incorrectly classified as $> p_t$. Type II error indicates that mean disease incidence was incorrectly classified as $< p_t$, where p_t is a critical value of disease incidence, such as a management action threshold for a fungicide application. Mean and median N were calculated from all 208 row-level data sets or 104 yard-level data sets. The parameters α and β were specified to control type I and type II error rates.

^b Data sets collected at the row level consisted of a single transect sample of $N = 100$ (Oregon) or 75 (Washington) plants within a row. Yard-level data sets consisted of two transect data sets per yard. At the yard level, a correct decision indicates that correct decisions were made in both transects at the row level. Type I error indicates that at least one type I error was made at the row level in a yard. Type II error indicates that a type II error was made at the row level in a yard. Type I and II errors indicate that a type I error was made in one row-level data set, and a type II error was made in the other row-level data in that yard.

^c $p_t = (p_0 + p_1)/2$; $p_0 = 0.0125$ and $p_1 = 0.0375$. For sequential classification, p_t is defined as some critical value of disease incidence, such as a management action threshold for a fungicide application, and p_0 and p_1 represent the lower and upper boundaries of disease incidence, such that when the true incidence of disease, p , is equal to or less than p_0 , the field is classified correctly at least $100(1 - \alpha)\%$ of the time, and when the true incidence of disease is equal to or greater than p_1 , the field is classified correctly at least $100(1 - \beta)\%$ of the time.

^d $p_t = (p_0 + p_1)/2$; $p_0 = 0.05$ and $p_1 = 0.15$.

^e Distribution and parameters estimated.

Achieving a C of 0.1 was not possible with most of the model validation data sets in which disease incidence was <0.2 because too few sampling units (150 or 200) were collected. However, achieving a C of 0.1 likely would require too much time and expense to be practical for routine disease assessments by growers and crop consultants. Using estimates of the parameters a and b from the binary power law (22, equation 17b) and a prespecified C of 0.1, 279 or 1,220 sampling units are required to estimate disease incidence when the true incidence is 0.05 or 0.01, respectively. Based on current knowledge of hop powdery mildew, designing a sampling strategy to attain $C = 0.1$ probably is not necessary if the objective is disease management (6,20). Additionally, because growers typically manage multiple hop yards, it is not practical to sample to attain $C = 0.1$ because it is more important to sample the greatest number of yards possible on a given farm than to obtain very precise estimates of disease from a few yards (36).

The sequential classification sampling models developed in this research allow for pest density to be classified as above or below various threshold values. Generally, smaller sample sizes are required when the objective of sampling is to reach a decision rather than to estimate disease incidence precisely. From a practical perspective, information on whether disease incidence is above or below a threshold value typically is sufficient for making most management decisions (4,27). The OC and ASN curves generated from bootstrap

simulations of the model construction data sets were very similar to the OC and ASN curves derived by Monte Carlo simulation. No more than 51, 33, 21, or 9 sampling units would need to be evaluated to classify disease incidence when $p_t = 0.025, 0.05, 0.10, \text{ or } 0.15$, respectively. This could dramatically reduce the time and expense associated with current hop powdery mildew sampling protocols that recommend sampling a minimum of 225 to 300 plants to estimate disease incidence for a yard (36). The number of sampling units needed to classify disease incidence could be reduced further by assuming a binomial distribution of disease incidence among leaves, which would reduce N , on average, by three to six plants depending on the error rate and p_t (Tables 1 and 2). However, it is questionable whether the modest decrease in N , perhaps a savings of 5 min of sampling, would justify the increase in type I and II error rates and loss of accuracy and precision.

Several factors affected the results of the sequential sampling plans developed in this research. The location in the yard where sampling was initiated affected the error rates of sequential classification. Disease incidence, on average, was greater in the second transect (closer to the center of the yard) than the first transect, and correct decisions rates varied depending on p_t and which transect was sampled first. Mean disease incidence can be variable among transects within a single yard (36) and, clearly, can affect conclusions (i.e., the estimate of p) obtained from a sequen-

tial sample. However, it is not practical to manage hop powdery mildew differently in areas of a yard where disease incidence may not be uniform. Control measures are applied to an entire yard if disease is above p_t in any area of the yard, because the risk of not treating a yard when a treatment is warranted is a more serious error than applying a treatment unnecessarily (25). Therefore, sampling should be initiated in areas of the yard where disease incidence typically is expected to be greatest, and the entire yard should be treated if disease exceeds p_t during sampling of any transect or area of the yard. Intentionally introducing bias into the sampling plan (i.e., targeted sampling) also may reduce sampling time and costs (4). Achieving a nonbiased sample of leaves from plants also may be difficult because powdery mildew lesions on adaxial leaf surfaces are readily apparent, although overestimation of disease incidence was not apparent during validation of the models by simulated sampling (Fig. 4; Tables 1 and 2). Biased selection of diseased leaves could reduce the rate of type II errors, which many growers consider more serious and costly than type I errors (25).

Heterogeneity in the model validation data sets was greater than for the model construction data sets (36). Estimates of the binary power law parameters used in the analyses were derived from yard-level assessments made in 2000 and 2001 (36). Consequently, these parameters underestimated disease heterogeneity in the validation data sets. It is unknown whether simi-

Table 2. Correct decision and error rates (%) and mean and median number of samples required for sequential classification of hop powdery mildew on leaves in relation to the location at which sampling was initiated in the hop yard

Rates ^d	Original sampling order ^a				Reversed sampling order ^a			
	$p_t = 0.025^b$		$p_t = 0.10^c$		$p_t = 0.025^b$		$p_t = 0.10^c$	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Binomial								
Correct decision	91.3	88.5	89.4	89.4	85.6	88.5	94.2	94.2
Type I error	1.0	1.0	4.8	4.8	3.8	1.0	2.9	2.9
Type II error	7.7	10.6	5.8	5.8	10.6	10.6	2.9	2.9
Mean N	12.3	9.3	3.7	3.6	13.0	9.3	3.9	3.7
Median N	12.0	9.0	3.0	3.0	12.0	9.0	3.0	3.0
Beta-Binomial								
Correct decision	94.2	90.4	90.4	90.4	86.5	90.4	97.1	96.2
Type I error	1.0	1.0	4.8	4.8	3.8	1.0	2.9	2.9
Type II error	5.8	8.7	6.7	6.7	10.6	8.7	1.9	2.9
Mean N	17.2	12.2	7.1	5.7	18.1	12.2	7.1	5.8
Median N	16.0	12.0	5.0	4.0	16.0	12.0	5.0	4.0

^a In the original sampling order, sampling began in transect 1 and continued into transect 2 if additional sampling was necessary to classify disease incidence as above or below p_t at the specified error rates, where p_t = a critical value of disease incidence, such as a management action threshold for a fungicide application. In the reversed sampling order, transect 2 of the data set was sampled first, and sampling continued into transect 1. Data sets were collected by assessing $n = 10$ leaves from the first $N = 100$ (Oregon) or 75 (Washington) plants from each of two transects (rows) in the hop yard. The parameters α and β were specified to control type I and type II error rates.

^b $p_t = (p_0 + p_1)/2$; $p_0 = 0.0125$ and $p_1 = 0.0375$. For sequential classification, p_t is defined as some critical value of disease incidence, such as a management action threshold for a fungicide application, and p_0 and p_1 represent the lower and upper boundaries of disease incidence, such that when the true incidence of disease, p , is equal to or less than p_0 , the field is classified correctly at least $100(1 - \alpha)\%$ of the time, and when the true incidence of disease is equal to or greater than p_1 , the field is classified correctly at least $100(1 - \beta)\%$ of the time.

^c $p_t = (p_0 + p_1)/2$; $p_0 = 0.05$ and $p_1 = 0.15$.

^d Distribution and rates (α and β). A correct decision indicates that mean disease incidence was correctly classified above or below p_t for the yard. Type I error indicates that mean disease incidence was incorrectly classified as greater than p_t for the yard. Type II error indicates that mean disease incidence was incorrectly classified as less than p_t for the yard. Mean and median N were calculated from all 104 yard-level data sets. Results are presented for stop lines generated by Wald's sequential probability ratio test for the binomial and beta-binomial cases (34, equations A1 to A6).

lar results would be found if another set of validation data was collected; however, from a practical standpoint, slight underestimation had little impact on the outcome because prescribed error rates generally were achieved with the validation data. It is possible to merge data sets and estimate a new set of power law parameters, but this would leave no independent data sets for model validation.

An important barrier that may prevent implementation of sequential sampling is the lack of formal disease thresholds for initiating control measures for hop powdery mildew. Economic thresholds have been derived for numerous arthropod pests (3,12,28,30) and monocyclic plant diseases (2,5,7,8,15), but few economic thresholds have been developed or proposed for polycyclic diseases (6,22). Where disease thresholds have been developed, they generally have been used to initiate a fungicide program (38) or to time fungicide applications in pathosystems where only a single application is needed to prevent substantial crop loss (9,10). Deriving disease thresholds for polycyclic diseases such as hop powdery mildew is more difficult than with many monocyclic diseases because of the short generation time of the pathogen and complex interactions of host, pathogen, and environment that result in crop loss (25). Maloy (25) suggests that monitoring for polycyclic diseases is problematic because the latent period between infection and disease development allows many diseases to expand to the point where economic management is not possible before detection. This may be true for hop powdery mildew, but many growers and integrated pest management personnel do consider disease status of a hop yard when deciding if, or at what intensity, control measures should be applied (36). Many growers have derived provisional action thresholds based on their past experience, the particular cultivar, and risk aversion (D. H. Gent and W. F. Mahaffee, unpublished). Thus, sequential classification plans for four values of p_i were developed to allow growers to select a disease threshold that reflects their provisional action threshold or to adjust the threshold based on infection risks (23,24).

This research demonstrated that it is possible to estimate or classify the incidence of powdery mildew on leaves of a hop yard over a wide range of disease levels by assessing relatively few sampling units. The sequential sampling plans developed in this study could be used to improve the efficiency and accuracy of disease assessments for current provisional action thresholds, and implementation of sequential sampling plans for hop powdery mildew may be one component of an overall decision-support system that considers cultivar susceptibility, past and forecast weather conditions, and grower risk aversion (23,24,37).

ACKNOWLEDGMENTS

Financial support was provided by USDA-ARS CRIS 303-5358-22000-030-00D and 5358-21000-035-00 and grants from the Hop Research Council, Washington State Commission on Pesticide Registration, and the Washington Hop Commission. We thank the many individuals who assisted in disease assessments, the hop growers who provided access to their yards and numerous insights, and L. du Toit and C. Ocamb for their review of an earlier draft of the manuscript.

LITERATURE CITED

- Barth, H. J., Klinke, C., and Schmidt, C. 1994. The Hop Atlas. Joh. Barth and Sohn, Nuremberg, Germany.
- Bejarano-Alcazar, J., Melero-Vara, J. M., Blanco-Lopez, M. A., and Jimenez-Diaz, R. M. 1995. Influence of inoculum density of defoliating and nondefoliating pathotypes of *Verticillium dahliae* on epidemics of *Verticillium wilt* of cotton in southern Spain. *Phytopathology* 85:1474-1481.
- Binns, M. R., and Nyrop, J. P. 1992. Sampling insect populations for the purpose of IPM decision making. *Annu. Rev. Entomol.* 37:427-453.
- Binns, M. R., Nyrop, J. P., and van der Werf, W. 2000. Sampling and Monitoring in Crop Protection: The Theoretical Basis for Designing Practical Decision Guides. CABI Publishing, Oxon, UK.
- Burt, O. R., and Ferris, H. 1996. Sequential decision rules for managing nematodes with crop rotations. *J. Nematol.* 28:457-474.
- Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley and Sons, New York.
- Chellemi, D. O. 2002. Nonchemical management of soilborne pests in fresh market vegetable production systems. *Phytopathology* 92:1367-1372.
- Crow, W. T., Weingartner, D. P., McSorley, R., and Dickson, D. W. 2000. Damage function and economic threshold for *Belonolaimus longicaudatus* on potato. *J. Nematol.* 32:318-322.
- Dillard, H. R., and Seem, R. C. 1990. Use of an action threshold for common maize rust to reduce crop loss in sweet corn. *Phytopathology* 80:846-849.
- Gaunt, R. E., and Cole, M. J. 1992. Sequential sampling for wheat stripe rust management. *Crop Prot.* 11:138-140.
- Gent, D. H., Mahaffee, W. F., and Turechek, W. W. 2006. Spatial heterogeneity of the incidence of powdery mildew on hop cones. *Plant Dis.* 90:1433-1440.
- Hoffman, M. P., Nyrop, J. P., Kirkwyland, J. J., Riggs, D. M., Gilrein, D. O., and Moyer, D. D. 1996. A sequential sampling plan for use in scheduling control of Lepidopterous pests for fresh market sweet corn. *J. Econ. Entomol.* 89:386-395.
- Hughes, G., and Madden, L. V. 1992. Aggregation and incidence of disease. *Plant Pathol.* 41:657-60.
- Hughes, G., Madden, L. V., and Munkvold, G. P. 1996. Cluster sampling for disease incidence data. *Phytopathology* 86:132-137.
- Kim, D. G., and Ferris, H. 2002. Relationship between crop losses and initial population densities of *Meloidogyne arenaria* in winter-grown oriental melon in Korea. *J. Nematol.* 34:43-49.
- Lalancette, N., and Hickey, K. D. 1986. An apple powdery mildew model based on plant growth, primary inoculum, and fungicide concentration. *Phytopathology* 76:1176-1182.
- MacKinnon, D. 2006. USA Hops: 2005 Statistical Report. Hop Growers of America, Yakima, WA.
- Madden, L. V., and Hughes, G. 1994. BBD-computer software for fitting the beta-binomial distribution to disease incidence data. *Plant Dis.* 78:536-540.
- Madden, L. V., and Hughes, G. 1995. Plant disease incidence: distribution, heterogeneity, and temporal analysis. *Annu. Rev. Phytopathol.* 33:529-564.
- Madden, L. V., and Hughes, G. 1999. Sampling for plant disease incidence. *Phytopathology* 89:1088-1103.
- Madden, L. V., Hughes, G., and Ellis, M. A. 1995. Spatial heterogeneity of the incidence of grape downy mildew. *Phytopathology* 85:269-275.
- Madden, L. V., Hughes, G., and Munkvold, G. P. 1996. Plant disease incidence: Inverse sampling, sequential sampling, and confidence intervals when observed mean incidence is zero. *Crop Prot.* 15:621-632.
- Mahaffee, W. F., Thomas, C. S., Turechek, W. W., Ocamb, C. M., Nelson, M. E., Fox, A., and Gubler, W. D. 2003. Responding to an introduced pathogen: *Podosphaera macularis* (hop powdery mildew) in the Pacific Northwest. Online. *Plant Health Progress* doi:10.1094/PHP-2003-1113-07-RV.
- Mahaffee, W. F., Turechek, W. W., and Ocamb, C. M. 2003. Effect of variable temperature on infection severity of *Podosphaera macularis* on hops. *Phytopathology* 93:1587-1592.
- Maloy, O. C. 1993. Plant Disease Control: Principles and Practice. John Wiley and Sons, New York.
- Neve, R. A. 1991. Hops. Chapman and Hall, London.
- Nyrop, J. P., Binns, M. R., and van der Werf, W. 1999. Sampling for IPM decision making: where should we invest time and resources? *Phytopathology* 89:1104-1111.
- Nyrop, J. P., Villani, M. G., and Grant, J. A. 1995. Control decision rule for European chafer (Coleoptera: Scarabaeidae) larvae infesting turfgrass. *Environ. Entomol.* 24:521-528.
- Ocamb, C., Klein, R., Barbour, J., Griesbach, J., and Mahaffee, W. 1999. First report of hop powdery mildew in the Pacific Northwest. *Plant Dis.* 83:1072.
- Renkema, J. M., Nyrop, J. P., Difonzo, C., Sears, M. K., and Schaafsma, A. W. 2006. Control decision rule for European chafer (Coleoptera: Scarabaeidae) larvae in field corn. *J. Econ. Entomol.* 99:76-84.
- Royal, D. J. 1978. Powdery mildew of the hop. Pages 281-409 in: *The Powdery Mildews*. D. M. Spencer, ed. Academic Press, London.
- Ryan, B., Joiner, B., and Cryer, J. 2005. *Mini-tab Handbook*, 5th ed. Brooks/Cole, Belmont, CA.
- Taylor, L. R. 1961. Aggregation, variance, and the mean. *Nature* 189:732-735.
- Turechek, W. W., Ellis, M. A., and Madden, L. V. 2001. Sequential sampling for incidence of *Phomopsis* leaf blight of strawberry. *Phytopathology* 91:336-347.
- Turechek, W. W., and Madden, L. V. 1999. Spatial pattern analysis and sequential sampling for the incidence of leaf spot on strawberry in Ohio. *Plant Dis.* 83:992-1000.
- Turechek, W. W., and Mahaffee, W. F. 2004. Spatial pattern analysis of hop powdery mildew in the Pacific Northwest: Implications for sampling. *Phytopathology* 94:1116-1128.
- Turechek, W. W., Mahaffee, W. F., and Ocamb, C. M. 2001. Development of management strategies for hop powdery mildew in the Pacific Northwest. Online. *Plant Health Progress* doi:10.1094/PHP-2001-0313-01-RS.
- Vincellii, P. C., and Lorbeer, J. W. 1987. Sequential sampling plan for timing initial fungicide application to control *Botrytis* leaf blight of onion. *Phytopathology* 77:1301-1303.
- Wald, A. 1947. *Sequential Analysis*. John Wiley & Sons, New York.