

# Sequential Sampling for Estimation and Classification of the Incidence of Hop Powdery Mildew II: Cone 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 II: Cone sampling. *Plant Dis.* 91:1013-1020.

Sequential sampling models for estimation and classification of the incidence of powdery mildew (caused by *Podosphaera macularis*) on hop (*Humulus lupulus*) cones were developed using parameter estimates of the binary power law derived from the analysis of 221 transect data sets (model construction data set) collected from 41 hop yards sampled in Oregon and Washington from 2000 to 2005. Stop lines, models that determine when sufficient information has been collected to estimate mean disease incidence and stop sampling, for sequential estimation were validated by bootstrap simulation using a subset of 21 model construction data sets and simulated sampling of an additional 13 model construction data sets. Achieved coefficient of variation ( $C$ ) approached the prespecified  $C$  as the estimated disease incidence,  $\hat{p}$ , increased, although achieving a  $C$  of 0.1 was not possible for data sets in which  $\hat{p} < 0.03$  with the number of sampling units evaluated in this study. The 95% confidence interval of the median difference between  $\hat{p}$  of each yard (achieved by sequential sampling) and the true  $p$  of the original data set included 0 for all 21 data sets evaluated at levels of  $C$  of 0.1 and 0.2. For sequential classification, operating characteristic (OC) and average sample number (ASN) curves of the sequential sampling plans obtained by bootstrap analysis and simulated sampling were similar to the OC and ASN values determined by Monte Carlo simulation. Correct decisions of whether disease incidence was above or below prespecified thresholds ( $p_i$ ) were made for 84.6 or 100% of the data sets during simulated sampling when stop lines were determined assuming a binomial or beta-binomial distribution of disease incidence, respectively. However, the higher proportion of correct decisions obtained by assuming a beta-binomial distribution of disease incidence required, on average, sampling 3.9 more plants per sampling round to classify disease incidence compared with the binomial distribution. Use of these sequential sampling plans may aid growers in deciding the order in which to harvest hop yards to minimize the risk of a condition called “cone early maturity” caused by late-season infection of cones by *P. macularis*. Also, sequential sampling could aid in research efforts, such as efficacy trials, where many hop cones are assessed to determine disease incidence.

Powdery mildew (caused by *Podosphaera macularis*) of hop (*Humulus lupulus*) is an economically important disease in the Pacific Northwest, where nearly all U.S. commercial production of hop occurs (1). The pathogen may infect all above-

ground plant parts and crown buds (10), although crop losses result from reductions in cone yield and quality due to infection of female flowers (burs) and developing cones (18). Such infections can result in burr or cone abortion, severe browning, and distortion of cones when infection occurs early. Infection after cone elongation results in a condition referred to as “cone early maturity” (3,6). Diffuse infections, often not visible to the unaided eye, cause infected cones to ripen faster than healthy cones and often result in growers harvesting overripe hops. Overripe hops are prone to shattering or crumbling during picking and cleaning, which reduces yield. Overripe hops may become brown in color during drying operations and impart an earthy aroma to beer (3,6). Complete crop losses can result from cone early maturity in aroma cultivars due to buyer rejection of the crop and absence of alternative markets for diseased cones. In 2001 and 2002 in

Oregon, 50% of the contracted crop of the aroma cv. Willamette was rejected because of cone early maturity, resulting in \$5 million in losses (16). Observations in Oregon and Washington hop yards from 2001 to 2004 indicate that a 2- to 5-day delay in the harvest of infected cones can result in unacceptable quality of aroma hops, with subsequent rejection by the contracting brewer or marketing company (W. F. Mahaffee, unpublished).

Knowledge of powdery mildew incidence on hop cones could be used by growers to schedule harvest dates in order to minimize losses due to cone early maturity. Current sampling protocols for estimating powdery mildew incidence on cones are impractical to implement on a large scale, because a 1,500-cone sample from a yard takes approximately 5 to 7 h to collect and assess (D. H. Gent and W. F. Mahaffee, unpublished). Furthermore, this sample size may not be sufficient to estimate disease incidence in large yards because of spatial aggregation of the disease (4). Gent et al. (4) generated fixed sampling curves to aid in assessing disease incidence of cones in order to reduce sampling costs. However, the use of fixed sampling curves requires a reasonable initial estimate of disease in order to obtain a precise estimate of disease without oversampling (15). The development of sequential sampling curves for estimating or classifying the incidence of disease above or below some threshold value could reduce sampling time and expense (2). Rapid assessment of the disease status of cones also may facilitate efforts to derive economic action thresholds for management of the foliar phase of powdery mildew on hop, and guide decisions about harvest dates for yards because the incidence of powdery mildew on foliage and cones is positively correlated (23).

The objective of this research was to develop and evaluate sequential sampling models for estimating and classifying the incidence of powdery mildew on hop cones. Parameters of the binary power law (8,12,13) derived from spatial pattern analysis of the incidence of hop powdery mildew on cones (4) were used to develop the sampling models. This work complements the sequential sampling models developed for the leaf phase of hop pow-

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 27 March 2007.

doi:10.1094/PDIS-91-8-1013

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.

dery mildew in part I of this article series (5).

## MATERIALS AND METHODS

**Data.** The incidence of hop cones with powdery mildew was assessed from 221 transect data sets sampled from 41 commercial hop yards in Oregon and Washington from 2000 to 2005 (4) using a cluster sampling strategy (9). In each yard, cones were collected from lateral branches of an individual plant at heights of approximately 2.7, 3.7, and 5.5 m above the ground and bulked. A subset of 25 cones ( $n$ ) was selected arbitrarily from each plant and assessed for powdery mildew. An individual data set consisted of disease incidence of 25 cones taken from each of the first 60 plants ( $N$ ) along the transect, for a total of 1,500 cones per yard. Data from all transects where disease incidence,  $\hat{p}$ , was

$> 0$  were considered a model construction data set to generate sequential sampling models. A subset of 21 transects from the model construction data set where  $\hat{p} > 0$  (201 data sets) was selected to develop data sets used for bootstrap simulations, described below. These data sets were selected because they encompassed the range of  $\hat{p}$  and heterogeneity,  $\hat{\theta}$ , observed among the model construction data sets collected from 2000 to 2005. For the model construction data in which  $\hat{p} > 0$ , data sets were classified into seven disease incidence categories:  $0.01 < \hat{p} < 0.025$ ,  $0.025 \leq \hat{p} < 0.05$ ,  $0.05 \leq \hat{p} < 0.10$ ,  $0.10 \leq \hat{p} < 0.20$ ,  $0.20 \leq \hat{p} < 0.40$ ,  $0.40 \leq \hat{p} < 0.60$ , and  $\hat{p} \geq 0.60$ . Three data sets were selected randomly from each disease incidence class and represented 7 yards in Oregon (6 of cv. Willamette and 1 of cv. Sterling) and 14 yards in Washington (all

of cvs. Columbus, Tomahawk, and Zeus, which are genetically indistinguishable and collectively referred to as CTZ). A subset of 13 transects that encompassed the range of  $\hat{p}$  and  $\hat{\theta}$  observed in the model construction data set were selected arbitrarily to validate the sequential estimation or classification procedures by simulated sampling, as described below. The simulated sampling validation data sets consisted of data sets collected in each year from 2000 to 2005, representing five yards in Oregon (four of cv. Willamette and one of cv. Glacier) and eight yards in Washington (all of cv. CTZ). The range of estimated disease incidence,  $\hat{p}$ , in these yards was 0.0006 to 0.918, and the range of the beta-binomial heterogeneity parameter,  $\hat{\theta}$ , was 0 to 0.386 (Table 1), as determined by the software BBD (11).

Covariance analysis (4) on the same data sets indicated that the effect of hop cultivar and geographic location of the yard did not affect the slope or intercept parameters of the binary power law (8,12) in the model construction data set. The year of sampling, however, did affect parameter estimates, with a tendency for the intercept,  $\ln(A_x)$ , and slope,  $b$ , to increase with time. It is unclear whether the variability in the parameters was the result of natural variability about the "true" values of  $\ln(A_x)$  and  $b$ , or if  $\ln(A_x)$  and  $b$  were approaching their true values. The parameters of  $a$  and  $b$  from Taylor's power law (20) generally are regarded as robust across space and time, although the portability of parameter estimates of the binary form of the power law is unclear and parameter estimates may vary over time (4). Therefore, the former was assumed and estimates of  $\ln(A_x) = 1.81$  and  $b = 1.22$  were obtained by pooling data over years for developing the sequential sampling models described below.

**Sequential estimation.** Methods for development of sampling curves for sequential estimation are fully described in part I of this article series (5) and by Turechek et al. (21). A summary of the methods is presented in this study. Sequential estimation stop lines were generated using a Mathcad version 13 (Mathsoft Inc., Cambridge, MA) worksheet (21) to solve equation 5 presented by Gent et al. (5) for  $T_N$  when  $N = 1$  to 500 and  $C = 0.1$  or  $0.2$ , with parameter estimates of the binary power law obtained as described above. Evaluation of the sequential estimation sampling plan was conducted by bootstrap simulation using the bootstrap-simulation data set.

Bootstrap simulation was conducted using a macro of Minitab version 14 (Minitab Inc., State College, PA) as described previously (5,21). For a single bootstrap simulation, sampling units ( $n = 25$  cones per plant) were sampled randomly one at a time with replacement from a transect, and the number of diseased cones ( $T_N$ ) was counted and summed. Sampling ceased

**Table 1.** Disease incidence ( $\hat{p}$ ), heterogeneity parameter of the beta-binomial distribution ( $\hat{\theta}$ ), and the index of dispersion ( $D$ ) of data sets used for bootstrap analysis and simulated sampling evaluation of sequential sampling models for hop powdery mildew<sup>a</sup>

Powdery mildew data set <sup>b</sup>			Measures of aggregation		
Year	Location	Cultivar	$\hat{p}$	$\hat{\theta}$	$D$
Bootstrap-simulation					
2004	Oregon	Willamette	0.01	0.27	1.65
2001	Oregon	Sterling	0.02	0.007	1.18
2004	Oregon	Willamette	0.02	0.062	2.44
2003	Oregon	Willamette	0.03	0.097	3.18
2003	Washington	CTZ	0.05	0.038	1.91
2004	Washington	CTZ	0.05	0.021	1.53
2000	Washington	CTZ	0.05	0.006	1.16
2000	Washington	CTZ	0.07	0.025	1.62
2002	Oregon	Willamette	0.07	0.027	1.66
2004	Oregon	Willamette	0.11	0.050	2.17
2005	Washington	CTZ	0.19	0.098	3.20
2004	Oregon	Willamette	0.20	0.11	3.35
2002	Washington	CTZ	0.22	0.11	3.40
2002	Washington	CTZ	0.32	0.13	3.80
2005	Washington	CTZ	0.37	0.15	4.19
2001	Washington	CTZ	0.40	0.091	3.06
2001	Washington	CTZ	0.49	0.022	1.54
2002	Washington	CTZ	0.54	0.37	7.78
2000	Washington	CTZ	0.62	0.16	4.40
2005	Washington	CTZ	0.74	0.24	5.68
2005	Washington	CTZ	0.92	0.13	3.85
Simulated sampling validation <sup>c</sup>					
2000	Washington	CTZ	0.0007	0	1.00
2005	Oregon	Willamette	0.002	0	0.97
2002	Oregon	Willamette	0.01	0.03	1.65
2003	Oregon	Willamette	0.03	0.10	3.18
2003	Washington	CTZ	0.05	0.04	1.91
2004	Washington	CTZ	0.05	0.02	1.53
2000	Washington	CTZ	0.07	0.08	2.76
2002	Oregon	Willamette	0.07	0.03	1.66
2004	Oregon	Glacier	0.11	0.05	2.17
2005	Washington	CTZ	0.19	0.10	3.20
2001	Washington	CTZ	0.44	0.09	3.06
2002	Washington	CTZ	0.55	0.39	7.78
2005	Washington	CTZ	0.92	0.13	3.85

<sup>a</sup> The parameter  $\theta$  of the beta-binomial distribution provides a measure of variation in disease incidence ( $\hat{p}$ ) per sampling unit, referred to as heterogeneity (12). The index of dispersion,  $D$ , is the ratio of the observed variance of the incidence of disease among sampling units to the expected variance under a binomial distribution (12).  $\hat{\theta}$  and  $D$  were calculated using the software BBD (11).

<sup>b</sup> A data set consisted of a single transect through a hop yard, where the incidence of hop powdery mildew was determined from  $n = 25$  cones sampled from each of the first  $N = 60$  plants along the transect (4).

<sup>c</sup> Slope and intercept parameters of the binary form of Taylor's power law (20) for the simulated sampling data sets were  $\ln(A_x) = 0.84$  and  $b = 1.26$ , respectively ( $R^2 = 0.98$ ), and did not vary significantly from the model construction data set ( $P = 0.3132$  and  $0.5967$ , respectively).

when the cumulative number of diseased cones exceeded the model  $T_N$ , or when the number of sampling units equaled the total number of sampling units in the data set (60 plants). A minimum of three sampling units was collected. The bootstrap simulation was conducted 100 times for each data set and specified value of  $C$ . The achieved  $C$  and moment estimate of  $p$  were calculated for each simulation, and the results were summarized in box plots as described previously (5,21).

The sequential estimation procedure was evaluated further by simulated sequential sampling of the 13 simulated sampling validation data sets using a Minitab macro as described previously (5,21). For a given data set, the data for the sampling units were entered into the macro in the same order they were collected in the field. The macro simulated sampling of diseased cones collected from the sampling units, and tallied the cumulative number of diseased cones until the cumulative number of disease cones exceeded the model  $T_N$  calculated from equation 5 in Gent et al. (5), or all sampling units were sampled. Estimates of  $\hat{p}$ , achieved  $C$ , and achieved  $N$  were calculated and related to the original (true)  $p$  of the data set using scatter plots of the difference between  $p$  and the achieved  $\hat{p}$ ,  $C$ , and  $N$ .

**Sequential classification.** Economic thresholds or economic injury levels for hop powdery mildew on cones have not been developed empirically for aroma or bittering cultivars. Therefore, sequential classification plans for powdery mildew on hop cones were developed at four incidence thresholds ( $p_i$ ) for initial evaluation:  $p_i = 0.05$  ( $p_0 = 0.025$ ,  $p_1 = 0.075$ );  $p_i = 0.10$  ( $p_0 = 0.05$ ,  $p_1 = 0.15$ );  $p_i = 0.15$  ( $p_0 = 0.05$ ,  $p_1 = 0.25$ ); and  $p_i = 0.20$  ( $p_0 = 0.10$ ,  $p_1 = 0.30$ ), where  $p_0$  and  $p_1$  represent the lower and upper boundaries of disease incidence such that when the true incidence of disease,  $p$ , is  $\leq p_0$ , the field is classified correctly at least  $100(1 - \alpha)\%$  of the time; and when the true incidence of disease is  $\geq p_1$ , 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  $H_0: p \leq p_i$  against the alternative hypothesis  $H_1: p > p_i$ , respectively. These threshold values were evaluated at each of four combinations of  $\alpha$  (0.05 and 0.10) and  $\beta$  (0.05 and 0.10). Thus, 16 pairs of stop lines were developed and evaluated by visual inspection of plots of their operating characteristic (OC) and average sample number (ASN) curves derived by Monte-Carlo simulations (5,21).

Four pairs of stop lines were selected for further evaluation by bootstrap simulation and simulated sampling. The parameters of the selected stop lines were:  $p_i = 0.05$  ( $p_0 = 0.025$ ,  $p_1 = 0.075$ ) and  $p_i = 0.15$  ( $p_0 = 0.05$ ,  $p_1 = 0.25$ ) at each of two error rates ( $\alpha = \beta = 0.05$  and  $\alpha = \beta = 0.10$ ). For a given boot-

strap simulation, sampling units ( $n = 25$  cones per plant) were sampled randomly one at a time, with replacement, from among all sampling units from the data set. Sampling ceased when the cumulative number of diseased cones,  $T_N$ , exceeded either 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 that a representative sample was collected. The achieved OC and ASN then were calculated (5). The bootstrap simulation was conducted 100 times for each data set and specified values of  $p_i$ ,  $p_0$ ,  $p_1$ ,  $\alpha$ , and  $\beta$ .

The four selected sequential classification sampling plans were evaluated by simulated sampling as described above for sequential estimation, but with stop lines calculated based on binomial and beta-binomial distributions (21, equations A1 to A6). The macro simulated sampling of diseased cones by summing the number of diseased cones in sequential sampling units, beginning with the first sampling unit, until the cumulative number of diseased cones exceeded the upper or lower stop lines of the model and a decision could be made as to whether the incidence of disease was above or below  $p_i$ , or the data set was sampled fully. To determine whether a correct decision was made,  $p$  calculated for the entire transect was assumed to represent the true  $p$ . The value of  $p$  was compared with  $p_i$  to determine the correct decision for that data set. The decision based on the results of simulated sequential classification then was compared with the correct decision. A type I or type II error was recorded if mean disease incidence was incorrectly classified as greater or less than  $p_i$ , respectively.

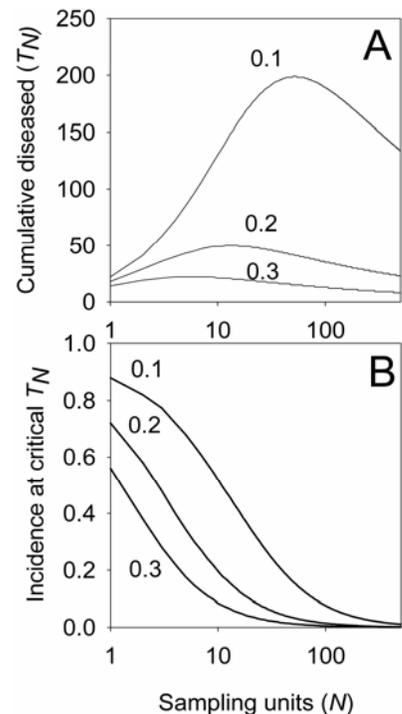
## RESULTS

**Spatial pattern analyses.** For the model construction data sets, disease incidence ( $\hat{p}$ ) ranged from 0 to 0.912 (median of 0.017) among yards. The log likelihood ratio test was significant for 92 (46%) data sets where  $\hat{p} > 0$ , indicating that the beta-binomial distribution provided a better fit to the data than the binomial distribution for these data sets. The heterogeneity parameter  $\hat{\theta}$  of the beta-binomial distribution was right-skewed and ranged from 0.01 to 0.39, with a median of 0.015. The slope and intercept parameters of the binary power law pooled over years were  $b = 1.22$  and  $\ln(A_i) = 1.81$ , respectively, and were significantly greater than 1 ( $P < 0.05$ ) based on  $t$  tests.

For the bootstrap-simulation data sets,  $\hat{p}$  ranged from 0.01 to 0.92, with a median of 0.019, and  $\hat{\theta}$  ranged from 0.006 to 0.386, with a median of 0.0497 (Table 1). The index of dispersion,  $D$ , ranged from 1.16 to 7.78, with a median of 3.06. For the simulated sampling validation data sets,  $\hat{p}$ ,  $\hat{\theta}$ , and  $D$  ranged from 0.0007 to 0.92 (median of 0.065), 0 to 0.37 (median

of 0.05), and 0.97 to 7.78 (median of 2.17), respectively (Table 1). The binary power law provided an excellent fit to the simulated sampling validation data sets ( $R^2 = 0.98$ ), with intercept and slope parameter estimates of  $\ln(A_i) = 0.84$  and  $b = 1.26$ , respectively. Based on  $t$  tests, the intercept and slope of the simulated sampling validation data sets were not significantly different from the model construction data sets ( $P = 0.31$  and 0.60, respectively).

**Sequential estimation.** Bootstrap simulation. Sequential estimation stop limits are presented in Figure 1. As expected, precision ( $C$ ) increased as disease incidence decreased, and the number of diseased cones that needed to be sampled in order to end the sequential sampling increased. For  $C = 0.2$ , confidence intervals based on the sign test included, or were less than, 0.2 for 11 data sets (61.1%) when  $\hat{p} \geq 0.025$  (Fig. 2A). It was not possible to achieve  $C = 0.2$  for  $0.01 \leq \hat{p} \leq 0.025$  (data sets 1 to 3), even when the sample size equaled the total number of sampling units (60 plants) in the data set (Fig. 2E). The achieved  $C$  tended to decrease with increasing  $\hat{p}$ , and  $C$  was less than 0.2 for 9 (42.9%) of the data sets. For  $C = 0.1$ , the median achieved  $C$  was



**Fig. 1.** Sequential estimation stop limits for estimating the mean incidence of powdery mildew on hop cones with a coefficient of variation ( $C$ ) of 0.1, 0.2, and 0.3. **A**, Cumulative number of diseased cones ( $T_N$ ) versus the total number of sampling units ( $N$ ) according to equation 5 in Gent et al. (5), with binary power law parameters  $b = 1.22$  and  $a = 0.15$ . **B**, Mean disease incidence at critical  $T_N$  (stop limit) in relation to  $N$ . Sampling ceases when the cumulative number of diseased cones crossed the critical  $T_N$  in A; mean disease incidence was then calculated as  $T_N/nN$ .

greater than the prespecified value in 16 (76.2%) of the data sets, although the median achieved  $C$  was close to 0.1 (i.e., within 0.07) for the 17 data sets for which  $\hat{p} > 0.03$ .

Median values of the difference between  $p$  and  $\hat{p}$  included 0 for the 21 data sets evaluated at  $C = 0.2$  and 0.1 (Fig. 2C and D). When  $p$  was low ( $<0.10$ ),  $p \cong \hat{p}$  because the data sets were fully sampled (Fig. 2E and F). Variation in the difference between  $p$  and  $\hat{p}$  tended to increase with increasing  $p$  (5,21). The median number of sample units required to estimate  $p$  decreased as  $p$  increased, and was greater for  $C = 0.1$  than 0.2 (Fig. 2E and F). It was possible to estimate  $p$  by sampling the

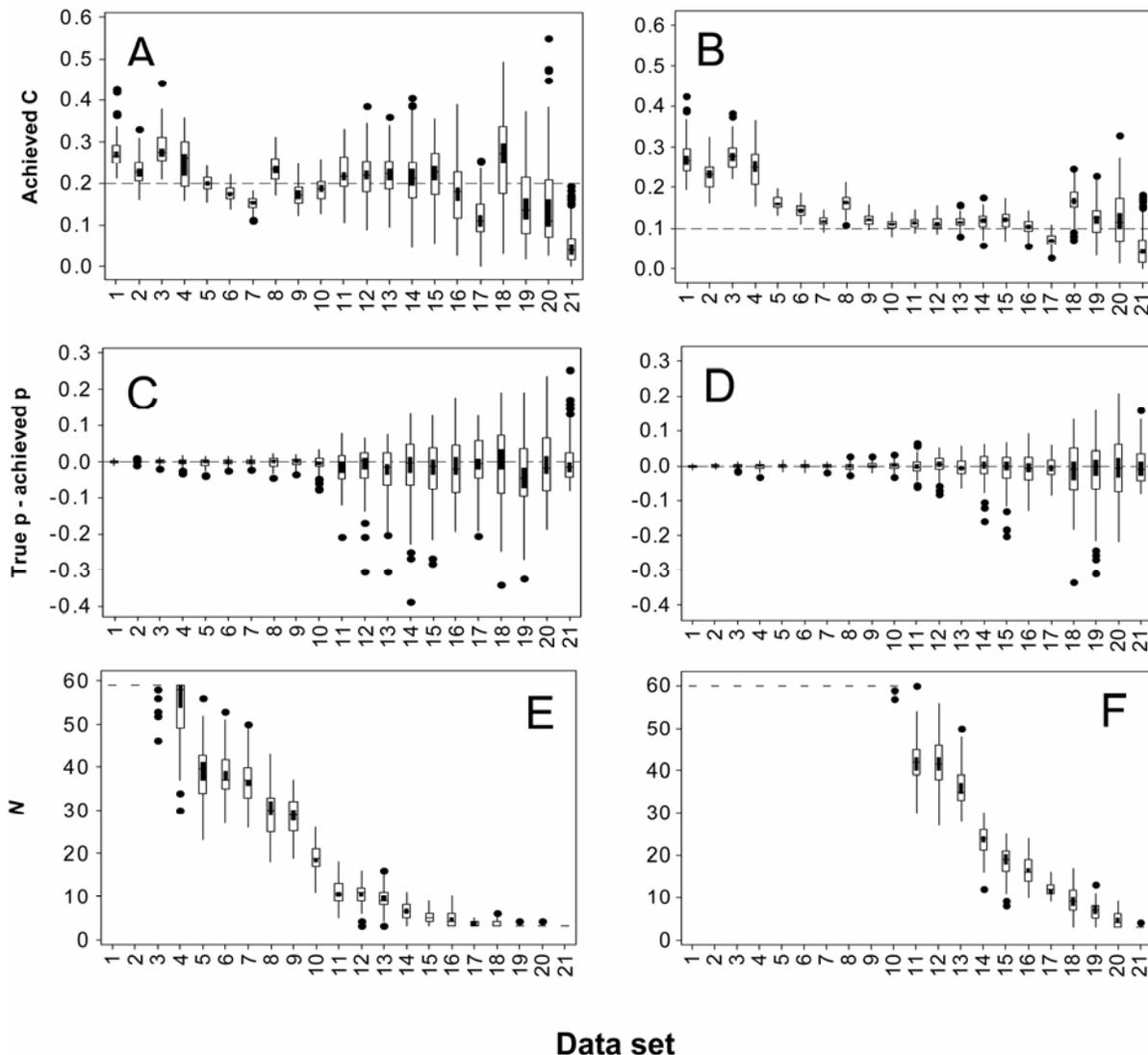
minimum number of sampling units specified (three plants) in all but two simulations when  $p > 0.6$  and  $C = 0.2$ .

**Simulated sampling validation.** For  $C = 0.2$ , the median achieved  $C$  was 0.2108 ( $Q_1 = 0.1863$  and  $Q_3 = 0.2755$ , where  $Q_1$  and  $Q_3$  are the first and third quartiles, respectively; Fig. 3A). Achieved  $C$  was less than the prespecified value except for data sets in which  $p < 0.025$ . For  $C = 0.1$ , achieving the specified  $C$  was not possible for data sets with  $p < 0.06$ , although the achieved  $C$  was near the specified values ( $<0.05$  units difference) when  $p$  was  $> 0.06$ . Differences between  $p$  and  $\hat{p}$  were similar, and never deviated more than 15% from  $p$  (Fig. 3B). For  $C = 0.2$  and  $C = 0.1$ , median values of

the difference between  $p$  and  $\hat{p}$  were 0, with  $Q_1 = -0.0018$  and  $Q_3 = 0.0086$ , and  $Q_1 = 0.0000$  and  $Q_3 = 0.0000$ , respectively. The difference between  $p$  and  $\hat{p}$  tended to increase with  $p$ , as observed by bootstrap simulation (Fig. 2C and D).

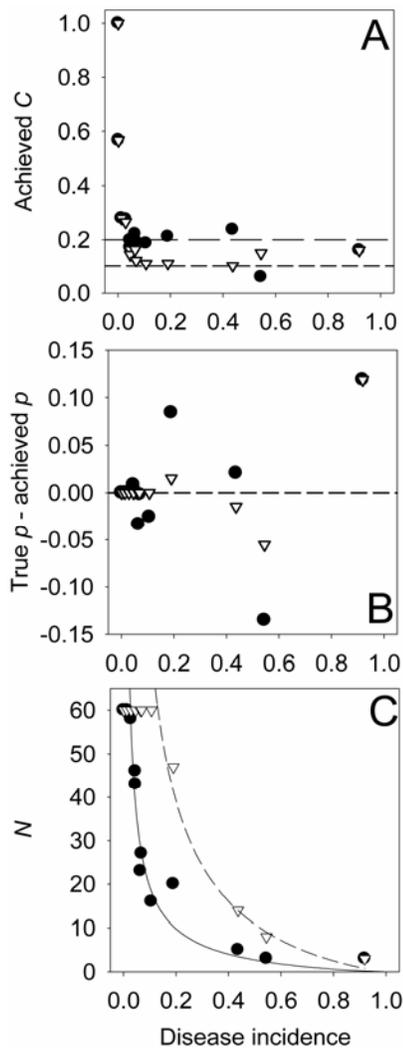
Sample size attained by sequential estimation was inversely related to  $p$  and similar to the expected sample sizes predicted (13, equation 17b) using  $a$  and  $b$  calculated from the model construction data sets. Data sets were fully sampled when  $C = 0.1$  and  $p < 0.15$ , and when  $C = 0.2$  and  $p < 0.025$  (Fig. 3C).

**Sequential classification.** *Stop limit selection.* Pairs of stop lines for sequential classification of the incidence of diseased



**Fig. 2.** A and B, Box plots of the achieved coefficient of variation ( $C$ ), C and D, difference between the true incidence of hop cones with powdery mildew,  $p$ , and the estimated  $p$  from the sequential sample ( $\hat{p}$ ), and E and F, achieved sample size ( $N$ ) for estimating the incidence of powdery mildew on hop cones from 100 simulated samplings of 21 bootstrap-simulation data sets used for bootstrap analysis. Preselected values of  $C$  were A, C, and E, 0.2 and B, D, and F, 0.1. Dashed lines in A and B indicate preselected values of  $C$ , and dashed lines in C and D indicate true  $p - \text{achieved } p = 0$ . Dashed lines are shown in E and F in place of box plots to indicate those data sets for which the preselected value of  $C$  could not be achieved even when the sample size equaled the total number of sampling units in the data set (i.e., 60). Sequential estimation stop lines were generated according to equation 2 of Gent et al. (5) with binary power law parameters of  $a = 0.15$  and  $b = 1.22$  (4). The 21 bootstrap-simulation data sets were chosen by selecting three data sets randomly from each of seven disease incidence classes ( $0.01 < p < 0.025$ ,  $0.025 < p < 0.05$ ,  $0.05 < p < 0.10$ ,  $0.10 < p < 0.20$ ,  $0.20 < p < 0.40$ ,  $0.40 < p < 0.60$ , and  $p > 0.60$ ) from among 221 model construction data sets collected from 2000 to 2005 in hop yards in Oregon and Washington. The 21 bootstrap-simulation data sets were arranged in ascending order of  $p$ . Box plots show the median (line in open boxes), middle 50% of the data (open box), confidence interval for the median based on the nonparametric sign-test (solid bar inside box), extremes of the data points (whiskers), and outliers (solid circles).

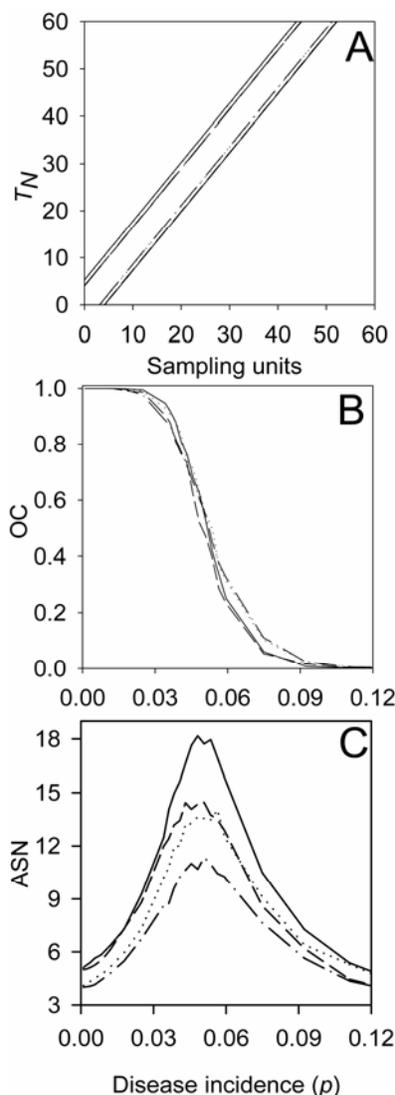
cones, and OC and ASN curves were derived by Monte Carlo simulation (Fig. 4A to C) for four combinations of  $\alpha$  and  $\beta$  error rates for  $p_t = 0.05$ . For  $\alpha = \beta = 0.05$ , varying  $p_t$  as well as the difference between  $p_0$  and  $p_1$  had a large effect on the slope and area between the stop lines, respectively (Fig. 5A). Steepness of the OC curve and height of the ASN curve decreased as the magnitude of the difference between  $p_0$  and  $p_1$  increased (Fig. 5B and C). For a given combination of  $\alpha = \beta$ , the



**Fig. 3.** Results of sequential sampling for estimation validation for 13 representative data sets collected from hop yards in Oregon and Washington during 2000 to 2005. **A**, Relationship between true disease incidence ( $p$ ) and the achieved coefficient of variation ( $C$ ). **B**, Relationship between  $p$  and estimated disease incidence ( $\hat{p}$ ) from the sequential sample. **C**, Relationship between  $p$  and number of sampling units ( $N$ ) collected. Solid circles and open triangles are data sets with preselected values of  $C = 0.2$  or  $0.1$ , respectively. The dashed lines in A are the prespecified values of  $C = 0.2$  (long dash) and  $0.1$  (short dash). The long-dashed line in B is set at  $(p - \hat{p}) = 0$ . The solid and dashed lines in C are the expected number of sampling units given the true disease incidence of the data set for  $C = 0.2$  and  $0.1$ , respectively, as calculated by Madden and Hughes (13, equation 17b).

distance between  $p_0$  and  $p_1$  determined the height of the ASN curve (i.e., how quickly a decision could be made; Fig. 5A and C). The wider limits (i.e.,  $p_0 = 0.10$  and  $p_1 = 0.30$ ) resulted in less sampling compared with narrow stop limits (i.e.,  $p_0 = 0.025$  and  $p_1 = 0.075$ ), but at a cost of higher error rates.

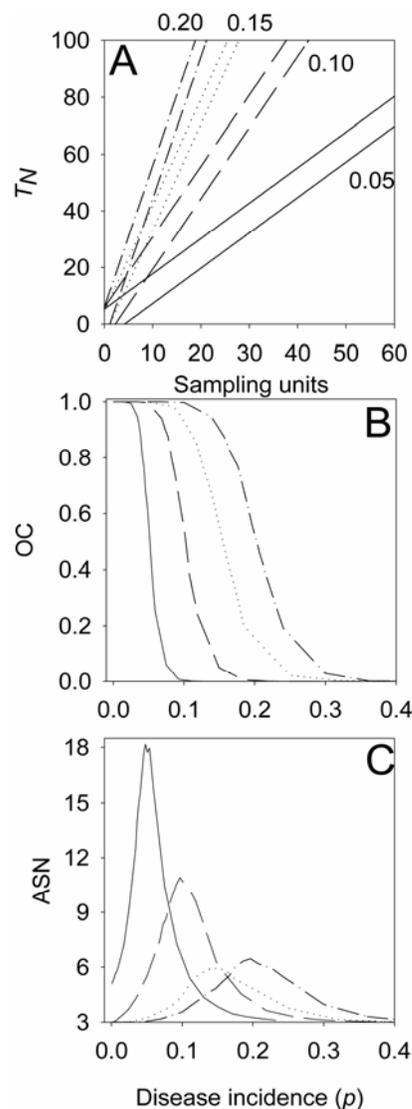
Increasing the error rates resulted in a slight decrease in the area between the stop



**Fig. 4.** Stop lines for classifying the incidence of powdery mildew on hop cones as above or below a critical threshold,  $p_t$ , from a cluster sample with  $n = 25$  cones per sampling unit. **A**, Stop lines based on Wald's sequence probability ratio test (25) with  $p_0 = 0.025$  and  $p_1 = 0.075$  ( $p_t = 0.05$ ) at four combinations of  $\alpha$  and  $\beta$ . **B**, Operating characteristic (OC) curves and **C**, average sample number (ASN) curves for sequential classification stop lines defined in A. Error probabilities are  $\alpha = 0.05$  and  $\beta = 0.05$  (solid line),  $\alpha = 0.10$  and  $\beta = 0.05$  (short 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. (7), with the parameter  $\theta$  as a function of disease incidence according to the binary power law (described in the text), where  $a = 0.15$  and  $b = 1.22$ .

lines (Fig. 4A) and had no effect on the slope of the stop lines, as predicted (21; equation A4). The steepness of the OC curve was changed little by decreasing  $\alpha$  or  $\beta$  (Fig. 4B); however, ASN increased by 10 sampling units when  $p$  was near  $p_t$  and when  $\alpha$  and  $\beta$  were controlled at 0.05 compared with 0.10 (Fig. 4C).

**Bootstrap simulation.** Based on results of the Monte Carlo simulation, four sampling plans were selected for further



**Fig. 5. 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 (25) for the incidence of powdery mildew on hop cones from a cluster sample with  $n = 25$  cones per sampling unit. OC and ASN curves were determined by 1,000 Monte Carlo simulations using a modification of the algorithm described by Hoffman et al. (7), with the parameter  $\theta$  described as a function of disease incidence (described in the text and by Gent et al. [4]), where  $a = 0.15$  and  $b = 1.22$ . 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.

evaluation by bootstrap analysis. The parameters of these sampling plans were  $p_t = 0.05$  and  $p_t = 0.15$  at error rates of  $\alpha = \beta = 0.05$  and  $\alpha = \beta = 0.10$  (Fig. 6A to D). Achieved OC and ASN among the 21 data sets were similar to the OC and ASN curves derived by Monte Carlo simulation, with two exceptions. ASN for the data sets where  $\hat{p} = 0.047$  and  $0.052$  were approximately 7 ( $\alpha = \beta = 0.05$ ) and 5 ( $\alpha = \beta = 0.10$ ) sampling units, respectively, greater than predicted by Monte Carlo simulation. The beta-binomial parameter  $\theta$  in these data sets was 0.021 ( $\hat{p} = 0.047$ ) and 0.21 ( $\hat{p} = 0.052$ ), although  $\theta$  predicted by the binary power law for these data sets was 0.0399 and 0.0692, respectively (4, Fig. 2)

**Simulated sampling validation.** Correct decisions were made for 12 (92%) data sets at  $p_t = 0.05$  and  $0.15$  with error rates of  $\alpha = \beta = 0.05$  or  $0.10$ , respectively, when stop lines were calculated assuming a binomial distribution of disease incidence (Table 2). The type I error occurred when  $p_t = 0.05$  for a data set with  $p = 0.047$ , and the type II error occurred when  $p_t = 0.15$

for a data set with  $p = 0.19$ . Type I and II errors were identical for stop lines calculated using a beta-binomial approximation, with the exception of  $p_t = 0.05$  with  $\alpha = \beta = 0.05$ . With these parameters, no classification errors were made for the 13 data sets evaluated. Median and mean  $N$  over the 13 data sets were greater by 2 to 3.93 sampling units, respectively, for the beta-binomial stop lines than the binomial stop lines for  $p_t = 0.05$  (Table 2). The greatest difference between  $N$  of the binomial and beta-binomial stop lines was observed for data sets with  $p < 0.02$  from  $p_t$  (Fig. 6C). For the data set where the type I error occurred at  $\alpha = \beta = 0.05$  using the binomial stop lines, 3 sampling units were assessed using the binomial stop lines and 26 were assessed using the beta-binomial stop lines. Mean and median  $N$  were identical for the binomial and beta-binomial stop lines at  $p_t = 0.15$ . Disease incidence was classified after assessing three sampling units for 11 (84.6%) of the data sets, and 4 sampling units for two data sets with  $\hat{p}$  close to  $p_t$  ( $\hat{p} = 0.11$  and  $0.19$ ).

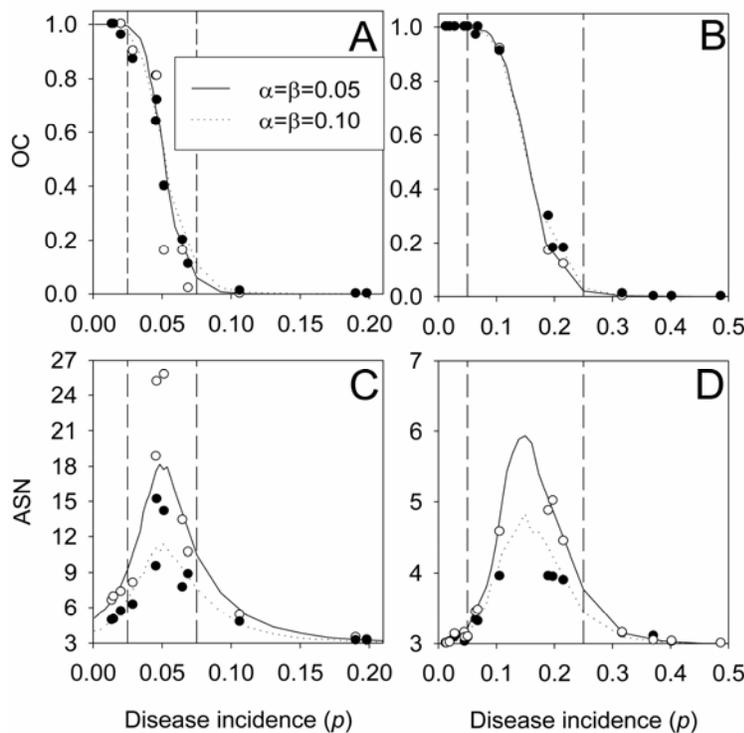
## DISCUSSION

Evaluations of hop cones for powdery mildew is labor and time intensive, and tends to be cost prohibitive for routine use in disease management. Therefore, hop growers in the Pacific Northwest typically do not make quantitative estimates of powdery mildew incidence in cones prior to harvest. The sequential sampling plans developed in this study for estimating and classifying the incidence of powdery mildew on cones performed well when evaluated by bootstrap analysis and simulated sampling of data sets with similar statistical properties. Implementation of these sampling plans should reduce sampling costs without compromising precision of disease assessments.

Unfortunately, independent data sets were not available to validate the sampling plans by simulated sampling. The use of resampling statistics is typical for many model validation studies (2,14,22) and, in some cases, sampling models are not validated at all (15). Binns et al. (2) note that the strength of resampling methods is that field-to-field variability of a sampling plan can be assessed. In this regard, the sampling plans evaluated in the current study performed as intended under simulated sampling with data sets that encompassed the range of disease incidence and heterogeneity typically observed in hop yards in Oregon and Washington from 2000 to 2005 (4).

Overall, the sequential estimation plan developed for cones performed similarly to the sampling plans for leaves (5). For both types of sampling plans, it was not possible (with the number of sampling units considered in this study) to estimate disease incidence with a desired precision of  $C = 0.1$  when disease incidence was low. This agrees with findings from other cropping systems, that very large sample sizes (e.g., >300 sampling units) are needed to estimate disease incidence with a  $C = 0.1$  (15,21,22). For both leaf (5) and cone estimates, the median differences between the true and estimated disease incidence were near 0, indicating that the sampling plans, on average, provided a close estimate of the true  $p$  of the data set. Apparently, variability in heterogeneity of the incidence of powdery mildew on hop cones (and leaves) among hop yards and over time had little effect on the performance of the sampling models. Further investigation is needed to determine the stability and portability of the binary power law parameters and practical implications for sampling.

For sequential classification, altering the distance between stop lines had a large influence on OC and ASN curves, whereas varying  $\alpha$  or  $\beta$  had relatively little effect on these curves. As noted by Nyrop et al. (17), the critical density ( $p_t$  in this study) of a pest is often the least-well-known parameter when sequential sampling plans



**Fig. 6.** 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 (25) for the incidence of powdery mildew on hop cones. Threshold values are A and C,  $p_0 = 0.025$ ,  $p_t = 0.075$ , and  $p_t = 0.05$ ; and B and D,  $p_0 = 0.05$ ,  $p_t = 0.25$ , and  $p_t = 0.15$ . OC and ASN curves were determined by 1,000 Monte Carlo simulations using a modification of the algorithm described by Hoffman et al. (7), with the parameter  $\theta$  as a function of mean disease incidence,  $p$ , according to the binary power law (described in the text and by Gent et al. [4]),  $a = 0.15$ , and  $b = 1.22$ . The achieved OC and ASN from 100 bootstrap simulations of sequential sampling for classification from 18 yard-level data sets; simulations were at  $\alpha = \beta = 0.05$  (open circles and dotted line) and at  $\alpha = \beta = 0.10$  (solid circles, solid line). The 21 bootstrap-simulation data sets were chosen by selecting three data sets randomly from each of seven disease incidence classes ( $0.01 < \hat{p} < 0.025$ ,  $0.025 < \hat{p} < 0.05$ ,  $0.05 < \hat{p} < 0.10$ ,  $0.10 < \hat{p} < 0.20$ ,  $0.20 < \hat{p} < 0.40$ ,  $0.40 < \hat{p} < 0.60$ , and  $\hat{p} > 0.60$ ) from among 201 model construction data sets collected from hop yards in Oregon and Washington from 2000 to 2005 (4). Data are not presented for data sets in which  $\hat{p} > 0.20$  when  $p_t = 0.05$ , or  $\hat{p} > 0.50$  when  $p_t = 0.15$  because OC = 0 and ASN  $\approx$  3 for these data sets.

for classification are developed, but have the greatest influence on the OC function and, therefore, correct decision rates. Unfortunately, economic thresholds or economic injury levels for powdery mildew on hop cones (or leaves) have not been developed. Data reported by Royal and Griffin (19) showed that powdery mildew incidence on hop cones from 2.3 to 17% was not significantly correlated with reductions in yield or  $\alpha$ -acid content of the cones in a fungicide trial conducted in England. Although this suggested a threshold for powdery mildew on cones of at least 17% for the cultivar evaluated in that study, Royal (18) cautioned that the absence of correlation between disease incidence and yield and  $\alpha$ -acid content may have resulted from late development of the disease or confounding interactions with hop downy mildew (caused by *Pseudoperonospora humuli*).

Mahaffee et al. (16) noted that, in 2001, 50% of the hop cone lots of the aroma cv. Willamette were rejected by the contracting brewery because of cone browning associated with diffuse infections by *Podosphaera macularis* that were largely nondetectable to the unaided eye. Apparently, thresholds for yield and quality losses in cones caused by powdery mildew differ between aroma hops and bittering hops, the latter intended for the  $\alpha$ -acid extract market in which cone appearance is of secondary importance to  $\alpha$ -acid content. As noted for powdery mildew on leaves (5), many growers likely have provisional

thresholds for powdery mildew of cones that are based on past experience, the intended market or contracting brewery of the crop, cultivar type (i.e., aroma or bittering cultivar), and their risk aversion. Like sequential classification plans for leaves (5), sampling plans were developed and evaluated for four values of  $p_t$  so that disease thresholds can be selected that are similar to the provisional threshold of individual growers or cultivar types. Although the thresholds are somewhat arbitrary, the thresholds are reasonable estimates and could improve disease management decision making compared with routine prophylactic treatments (2) or current sampling protocols (4). Uncertainties associated with the thresholds used in the sequential classification plans could be avoided if mean disease incidence was estimated by sequential estimation. Nonetheless, the sequential sampling models for cones should be helpful for improving the efficiency of assessing the incidence of powdery mildew in young cones and making decisions about late-season fungicide applications or determining harvest dates to minimize losses by cone early maturity (3).

Turechek et al. (23) reported that the mean incidence of powdery mildew on leaves explained 68% of the variability in the mean incidence of powdery mildew on cones. Potentially, the incidence of powdery mildew on leaves could be determined by sequential samplings and used to predict the incidence of diseased cones. These predictions could aid growers in

making management decisions for the foliar phase of the disease to limit the risk of severe cone infection. However, it is unclear when leaf assessments should be made to most accurately predict disease incidence on cones, or how well such assessments would perform under varying disease thresholds. Receiver operating characteristic curve analysis (24) may help to resolve these questions once predictive models for cone infection are developed. Linkage of sequential sampling models for the leaf and cone phase of the disease should become clearer after the relationship between the foliar and cone phases of hop powdery mildew are quantified more precisely.

In conclusion, the sequential sampling plans developed in this research were capable of estimating disease incidence with reasonable precision ( $C = 0.2$ ), or classifying disease incidence with acceptable accuracy, while assessing few sampling units. Sequential sampling overcomes the limitation of fixed sampling plans developed previously for powdery mildew on cones (4), and may aid in implementation of threshold-based management of hop powdery mildew on cones. More efficient sampling of powdery mildew incidence also may help growers to make better decisions regarding late season management of the disease on cones.

#### 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, and the Washington Hop Commission. We thank the many individuals who assisted in disease assessments, the hop growers who allowed us access to their yards, and L. du Toit and C. Ocam for reviewing an earlier draft of this manuscript.

#### LITERATURE CITED

- Barth, H. J., Klinke, C., and Schmidt, C. 1994. The Hop Atlas. Joh. Barth and Sohn, Nuremberg, Germany.
- 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.
- Coley-Smith, J. R. 1964. An association between so-called premature ripening of hops and the presence of powdery mildew (*Sphaerotheca macularis*). Rep. Dep. Hop Res. Wye Coll. For. 1963:30-1.
- 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.
- 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.
- Hammond, W. H. 1900. On red mould, or premature "going off" of hops. J. Southeast. Agric. Coll. Wye Engl. 9:19-20.
- 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.

**Table 2.** Correct decision and error rates (%), and the mean and median number of samples required for sequential sampling for classification of the incidence of powdery mildew on hop cones for stop lines generated by Wald's sequential probability ratio test for the binomial and beta-binomial distributions (21)

Distribution <sup>d</sup>	Critical incidence ( $p_t$ ) <sup>a</sup>			
	$p_t = 0.05^b$		$p_t = 0.15^c$	
	0.05	0.10	0.05	0.10
Binomial				
Correct decision	0.92	0.92	0.92	0.92
Type I error	0.08	0.08	0	0
Type II error	0	0	0.08	0.08
Mean $N$	3.15	3.08	3.15	3.15
Median $N$	3.00	3.00	3.00	3.00
Beta-Binomial				
Correct decision	1.00	0.92	0.92	0.92
Type I error	0	0.08	0	0
Type II error	0	0	0.08	0.08
Mean $N$	7.08	4.62	3.15	3.15
Median $N$	5.00	4.00	3.00	3.00

<sup>a</sup> For sequential classification,  $p_t$  is a critical value of disease incidence ( $p$ ) that is important for hypothesis testing, such as a management action threshold for a fungicide application. The parameters  $\alpha$  and  $\beta$  are pre-specified to control type I and type II error rates.

<sup>b</sup>  $p_t = (p_0 + p_1)/2$ ;  $p_0 = 0.025$  and  $p_1 = 0.075$ , where  $p_0$  and  $p_1$  are upper boundaries of disease incidence such that when  $p \leq p_0$ , the field is classified correctly at least  $100(1 - \alpha)\%$  of the time; and when  $p \geq p_1$ , the field is classified correctly at least  $100(1 - \beta)\%$  of the time.

<sup>c</sup>  $p_t = (p_0 + p_1)/2$ ;  $p_0 = 0.05$  and  $p_1 = 0.25$ .

<sup>d</sup> Distribution and error rates ( $\alpha = \beta$ ). A correct decision indicates that the mean disease incidence was correctly classified as above or below  $p_t$ . Type I and II errors occurred when mean disease incidence was incorrectly classified as greater than or less than  $p_t$  for a data set, respectively. Mean and median  $N$  were calculated from 13 model evaluation data sets. Data sets were collected by assessing  $n = 25$  cones from each of the first  $N = 60$  plants in a transect (row) in a hop yard in Oregon or Washington sampled between 2000 and 2005 (4).

8. Hughes, G., and Madden, L. V. 1992. Aggregation and incidence of disease. *Plant Pathol.* 41:657-660.
9. Hughes, G., Madden, L. V., and Munkvold, G. P. 1996. Cluster sampling for disease incidence data. *Phytopathology* 86:132-137.
10. Liyanage, A. de S., and Royal, D. J. 1976. Overwintering of *Sphaerotheca humuli*, the cause of hop powdery mildew. *Ann. Appl. Biol.* 83:381-394.
11. 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.
12. Madden, L. V., and Hughes, G. 1995. Plant disease incidence: distribution, heterogeneity, and temporal analysis. *Annu. Rev. Phytopathol.* 33:529-564.
13. Madden, L. V., and Hughes, G. 1999. Sampling for plant disease incidence. *Phytopathology* 89:1088-1113.
14. Madden, L. V., Hughes, G., and Ellis, M. A. 1995. Spatial heterogeneity of the incidence of grape downy mildew. *Phytopathology* 85:269-275.
15. 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.
16. 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.
17. 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.
18. Royal, D. J. 1978. Powdery mildew of the hop. Pages 281-409 in: *The Powdery Mildews*. D. M. Spencer, ed. Academic Press, London.
19. Royal, D. J., and Griffin, M. J. 1973. Side-effects of downy mildew fungicides on the incidence of hop powdery mildew (*Sphaerotheca humuli*). *Plant Pathol.* 22:129-133.
20. Taylor, L. R. 1961. Aggregation, variance, and the mean. *Nature* 189:732-735.
21. 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.
22. 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.
23. 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.
24. Turechek, W. W., and Wilcox, W. F. 2005. Evaluating predictors of apple scab with receiver operating characteristic curve analysis. *Phytopathology* 95:679-691.
25. Wald, A. 1947. *Sequential Analysis*. John Wiley & Sons, New York.