Enumerative and Binomial Sequential Sampling Plans for Soybean Aphid (Homoptera: Aphididae) in Soybean

E. W. HODGSON, E. C. BURKNESS, W. D. HUTCHISON, AND D. W. RAGSDALE

Department of Entomology, 219 Hodson Hall, University of Minnesota, St. Paul, MN 55108

J. Econ. Entomol. 97(6): 2127–2136 (2004)

ABSTRACT Since the discovery of the soybean aphid, *Aphis glycines* Matsumura, in midwestern U.S. soybean, *Glycine max* L., in 2000, the aphid has become a significant economic pest. Basic information about estimating population density within fields is unknown. Therefore, we developed two sampling plans to efficiently characterize *A. glycines* densities. Enumerative and binomial sequential plans were developed using 89 data sets collected from 10 commercial fields sampled during 2001–2003. Resampling software was used to validate the enumerative plan on whole plant counts, based on Taylor's power law parameters ($a = 9.157$ and $b = 1.543$). For research applications, the enumerative plan was modified to provide an actual precision level of 0.10 (SE/$\bar{x}$), which resulted in an average sample number of 310 individual plants. For integrated pest management (IPM) purposes, we developed an enumerative plan with an actual precision of 0.25, which resulted in an average sample number of 38 individual plants. For IPM applications, the binomial plan will likely be more practical. Binomial plans were developed using two tally thresholds at five action thresholds. Final analysis of the operating characteristic curve for each plan indicated that the tally threshold of ≥40 aphids per plant, and an action threshold of 0.837 (84% of the plants infested) provided the most correct treat (4%) and no-treat (95%) decisions, with very low incorrect treat (0.5%) and no-treat (0.5%) decisions. A tally threshold of ≥40 aphids per plant and action thresholds of 84% of plants infested is equivalent to a mean density of 250 aphids per plant, a recently recommended economic threshold. Using this threshold, the minimum required sample number for the binomial plan was 11 plants.

KEY WORDS *Aphis glycines*, enumerative sampling, binomial sampling, resampling validation

THE SOYBEAN APHID, *Aphis glycines* Matsumura, is an introduced pest from Asia and was discovered in North American soybean, *Glycine max* L., in 2000 (Ragsdale et al. 2004, Wu et al. 2004). When *A. glycines* invaded the United States and Canada, soybean fields were relatively open niches and establishment was rapid. Populations of *A. glycines* can now be found throughout the North Central states, particularly in Minnesota, Michigan, and Wisconsin (Venette and Ragsdale 2004). Since the initial discovery, *A. glycines* have colonized every soybean production county in Minnesota in <3 yr (MPR 2003). Yield loss due to aphid feeding or disease transmission now present significant challenges for soybean growers. Yield losses of 12–45% have been observed in strip trials (Ostlie 2001), and indirect damage from virus transmission is also possible (Hill et al. 2001). In 2003, >2.4 million ha of soybean was planted in Minnesota with >1.2 million ha treated with an insecticide at least once during the growing season (Landis et al. 2003). However, some insecticides may have been unnecessarily applied due to inaccurate estimates of *A. glycines* density.

Sampling methods for estimating within-field *A. glycines* densities are not clearly defined, and a common protocol is necessary before we can accurately monitor populations, evaluate economic thresholds, and assess treatment efficacy. Statistically sound sampling plans are needed for researchers, growers, crop consultants, and extension specialists to develop an integrated pest management (IPM) program for *A. glycines* in soybean. Sampling plans for *A. glycines* have been published in China, but they may not be applicable to populations in North America (Liu 1986, Huang et al. 1992, Su et al. 1996). Thoroughly developed sampling plans have been developed for other aphids in the United States, including the cabbage aphid, *Brevicoryne brassicae* (L.), and green peach aphid, *Myzus persicae* (Sulzer) (Wilson et al. 1983); the greenbug, *Schizaphis graminum* (Rondani) (Giles et al. 2000); and the bird cherry-oat aphid, *Rhopalosiphum padi* (L.) (Elliott et al. 2003). The methods for developing sampling plans in different systems can provide a framework for new pests such as *A. glycines.*

The purpose of designing a sampling protocol is to develop a guideline for collecting samples with an acceptable level of precision over a wide range of potential population densities (Nyrop and Binns 1991, Hutchison 1994). However, the desired precision is not necessarily achieved from one sampling session

0022-0493/04/2127–2136$04.00/0

(bout), but will be the expected precision for a number of sampling sessions (Hutchison et al. 1988). Recent entomological literature has demonstrated that a particular resampling software program, Resampling for Validation of Sample Plans (Naranjo and Hutchison 1997), has been effective for use with several insect pests in various crops worldwide, including *Acalymma vittatum* (F.) (Coleoptera: Chrysomelidae) in cucurbits (Burkness and Hutchison 1997, 1998); *Bemisa tabaci* (Gennadius) (Homoptera: Aleyrodidae) in cotton (Naranjo and Flint 1995); *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae) in eggplant (Hamilton et al. 1998); *Plutella xylostella* (L.) (Lepidoptera: Plutellidae) in broccoli and cauliflower (Hamilton et al. 2004); and *Pseudaletia unipuncta* (Haworth) (Lepidoptera: Noctuidae) in Azorean pastures (Silva et al. 2003). Naranjo and Hutchison (1997) demonstrated proper use of Resampling for Validation of Sample Plans with a comprehensive software review.

With only three field seasons of research since the establishment of *A. glycines* in North America, much of our knowledge about the dynamics of colonization and yield loss potential is incomplete. However, a tentative economic threshold of 250 aphids per plant through the pod set growth stage has been adopted for *A. glycines* for 2004 (NCSRP 2004). Estimating *A. glycines* densities with efficiency becomes especially important when environmental conditions promote exponential growth (i.e., making treatment decisions before the population exceeds the economic injury level). The growth potential for *A. glycines* can be dramatic under favorable conditions. McCornack et al. (2004) reports the optimal growth temperature for *A. glycines* is 27.8°C, at which populations can double in 1.5 d under a laboratory situation.

The purpose of this research was to develop and validate sequential sampling plans for both estimating *A. glycines* density (enumerative) and making pest management decisions (binomial) for research and IPM applications, respectively. Validating sampling plans using field-collected data sets and resampling software can also provide realistic mean densities compared with theoretical simulations (Hutchison 1994). Developing a sequential sampling plan with validation is crucial because of uncertainty inherent with any sampling procedure (Hutchison 1994, Naranjo and Hutchison 1997). The development of an efficient sampling plan to estimate *A. glycines* densities will reduce the time (cost) required for sampling commercial soybean and facilitate more accurate management decisions.

Materials and Methods

From 2001 to 2003, 10 commercial soybean fields were sampled for *A. glycines.* In total, 89 data sets were collected over 58 samples dates from central and southern Minnesota (Dakota, Houston, Olmsted, and Steele Counties). Fields ranged in size from 4.0 to 20.2 ha, and row spacing was at least 76 cm for all 10 fields. Nondestructive, whole plant counts were used to estimate aphid densities for all data sets. Of the 89 data sets, mean densities ranged from 3 to 2,773 *A. glycines* per plant, with seven data sets having a mean of at least 250 aphids per plant.

In each of the 10 fields, a systematic field design with a uniform grid pattern was used to estimate the spatial and temporal distribution. To ensure adequate coverage for each field at every collection date, the same points were sampled in the field at each collection date from early vegetative stages to seed set. The number of sample points varied between fields and was dependent on field size and shape (2001: field 1, 80; and field 2, 63; 2002: field 3, 60; field 4, 273; field 5, 298; and field 6, 80; and 2003: field 7, 174; field 8, 174; field 9, 160, and field 10, 160). Sample units were always collected within 2 m of each sample point. Sample units for fields 1 and 2 consisted of 10 whole plants selected at random. For fields 3–5 in 2002, each sample unit consisted of two whole plants. Field 6 in 2002 had variable sample units with whole plant counts, and fields 7–10 in 2003 had samples ranging between two and 10 whole plants. Aphid densities on each plant were recorded separately and ≈60 s spent estimating aphid density per plant with ≈30 s walking between sample points.

Enumerative Sampling. For resampling analysis, 74 of 89 independent data sets were randomly selected to calculate a and b values for Taylor's power law ($s^2 = am^b$), which describes the mean ($\bar{x}$)-to-variance (s^2) ratio and the dispersion pattern of a species (Taylor 1961). The a and b values from Taylor's power law were used to develop Green's enumerative sampling plan (Green 1970). Sampling plans for *A. glycines* can then be evaluated with resampling analysis. Fifteen data sets, representing a range of *A. glycines* densities (means ranged from 3 to 303 aphids per plant), were excluded from the original 89 for use in validating the sampling plan in Resampling for Validation of Sample Plans.

The sampling stop line for Green's plan is calculated by the following equation:

$$T_n \geq (an^{1-b}/(se/\bar{x})^2)^{1/(2-b)}$$

where T_n is the cumulative number of individuals sampled, n is the total number of samples, *se*/ is the precision desired by the user, and a and b are the Taylor's power law parameters. The simulation selected samples randomly from each of the 15 validation data sets until the stop line was exceeded. Resampling analysis for sample size was based on 500 iterative sampling runs. Two fixed levels of precision were selected for both intensive ecological (0.10) and pest management (0.25) purposes as suggested by Southwood (1978). Mean actual precision was calculated and compared with the desired precision of 0.10 and 0.25. Mean actual density, mean sample number, and mean actual precision values were obtained for each data set of 500 simulations. The average sample number from the 15 validation sets was then selected as the recommended sample size.

Binomial Sampling. For binomial analysis, all data sets were selected to create decision lines for the Wald

(1947) binomial sequential probability ratio test. Decision lines based on Wald's binomial plan were validated with Resampling for Validation of Sample Plans. The proportion of plants infested and operating characteristic (OC) values from the binomial simulations were used to determine which tally threshold and action threshold combination provided the highest probability of making a correct decision. Binomial sampling can be highly efficient compared with exhaustive counts, especially for pests such as *A. glycines*, where an actual population can exceed 2,000 aphids per plant. Incorporating a tally threshold (i.e., more than one insect per plant to be considered infested), can reduce sampling error and also minimize the sample size required for each field (Ward et al. 1985, 1986a, b).

Data sets were randomly selected and were iteratively resampled until a treat or no-treat decision was made. The test requires parameters for θ_1 and θ_2, the lower and upper boundaries, respectively, for the decision action threshold, and for α (type I) and β (type II) error rates (Jones 1994). All parameters were held constant for all tally threshold and action threshold combinations, including the upper and lower boundaries of the action threshold, θ_1 and θ_2 (= 0.10 above and below the action threshold, respectively), and α and β error, respectively. Stop lines were calculated using Wald's plan and were based on the optimum tally threshold and action threshold combination. These stop lines are defined as follows:

$$T_{n(t)} \geq Rx + Q \text{ and } T_{n(t)} \leq Rx - S$$

where $T_{n(t)}$ is the cumulative number of samples infested with at least t insects, and Q, R, and S are functions of α and β. From each simulation, the average proportion infested, average sample number, and the OC function are calculated and summarized. The OC function estimates the probability of not taking action against a pest population relative to the pest density (i.e., the proportion of plants infested) (Onsager 1976). In addition, resampling provides actual α and β values for each sample comparison.

We evaluated two tally thresholds of 20 and 40 aphids (i.e., ≥20 aphids or ≥40 aphids per plant to be considered infested, respectively). We considered these tally thresholds reasonable for consultants and cooperators for use in an IPM program. Other tally thresholds were considered (i.e., 50 and 100), but the R^2 values were not notably different from the lower thresholds (data not shown). Higher tally thresholds also diminish the convenience of creating a practical binomial sampling plan by increasing the minimum count of insects per plant. In addition to evaluating two tally thresholds, five mean densities were selected as potential action thresholds (i.e., 25, 50, 100, 150, and 250 aphids per plant). Mean densities examined represent a range of mean densities typical in commercial soybean and were proposed as tentative economic thresholds early in the establishment of *A. glycines*. Each action threshold was determined by examining the relationship between the mean density and proportion of plants infested for the data sets at a given tally threshold (20 or 40) and was derived from extrapolating values from the proportion of plants infested and mean curves (Fig. 1).

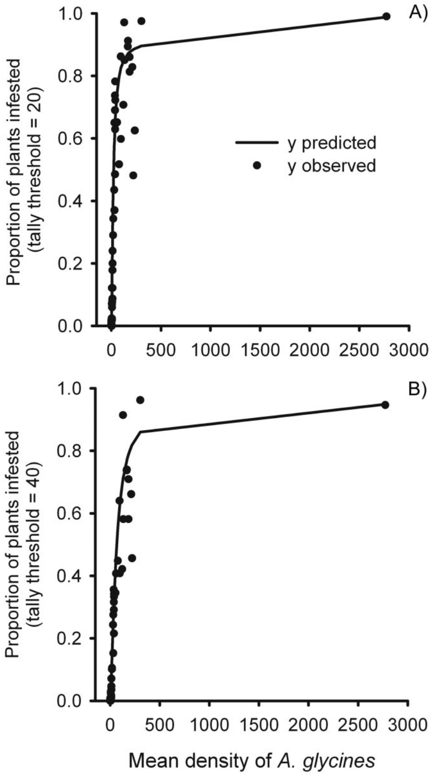

Fig. 1. Nonlinear relationship between the proportion of plants infested and the mean number of *A. glycines* per plant. The solid lines represent the fitted curve equation: $y^{-1} = 1.11 + (-3.45 * x) + (103.85/x^{1.5})$. (A) Tally threshold ≥20 aphids per plant, $R^2 = 0.9$. (B) Tally threshold ≥40 aphids per plant, $R^2 = 0.91$.

The probability of making a correct decision (i.e., correctly treating or correctly not treating) was calculated using a four-cell decision matrix described by Calvin et al. (1986) and Burkness et al. (1999). Cells A and D are correct decisions to treat (α) and not to treat (β), respectively, and cells B and C are incorrect decisions to treat and not to treat, respectively. For each data set evaluated, a proper decision is fixed by the magnitude of density and must be made as correct or incorrect decision in the matrix, where A + B = 1 or C + D = 1 (Burkness et al. 1999). Consequently, when the density is large enough to require a treatment, the probability of A = 1 − OC and the probability of B = OC; when the density is too low to require a treatment, the probability of C = 1 − OC and the probability of D = OC. The probability of making a correct decision (i.e., to treat or not to treat) was summarized for all data sets at each tally threshold with

$$1 = \Sigma p_i \, (A_i + D_i) + \Sigma p_i \, (B_i + C_i)$$

where p_i is the proportion of n data sets represented by data set i, A_i is the probability of making a correct treat decision, D_i is the probability of making a correct no-treat decision, B_i is the probability of making an incorrect no-treat decision, and C_i is the probability of making an incorrect treat decision. A suitable tally and action threshold combination should provide a greater probability of making correct decisions and minimize the probability of not treating when a pest is above the recommended economic threshold (Burkness et al. 1999).

Sampling Cost. Another way of analyzing enumerative and binomial sequential plans is to estimate the cost (in time) associated with sampling plants. Because the binomial sequential plan is not calculated with precision values, the two tally threshold sampling plans are not directly comparable. However, both plans have attributes that can be measured by the increasing cost of sampling time.

Fixed precision sampling plans can be directly compared by using relative net precision, which gives equal consideration to precision and time as variables (Pedigo et al. 1972). Higher relative net precision indicates a more efficient sampling plan and is calculated by the following:

$$\text{relative net precision} = (1/(RV * c)) * 100$$

where RV is the relative variation $(\text{se}/\bar{x}) * 100$ (Southwood 1978), and c is the total cost related to collecting the given number of samples usually measured in person-hours.

Binomial sequential plans are intuitively more time-efficient than enumerative sampling, because complete density estimates are not required. Although increasing tally thresholds (>1) will increase sampling time per unit, significant savings can still be realized compared with enumerative plans (Jones 1994). Furthermore, because higher tally thresholds provide more information per sample unit, fewer total sample units are typically required compared with enumerative sampling plans (Jones 1994).

To directly compare different binomial sampling plans, we developed a benefit-cost ratio. The benefit-cost ratio uses the proportion of total correct decisions, mentioned previously (Burkness et al. 1999), and the cost of obtaining density estimates (as per relative net precision; Pedigo et al. 1972). The benefit-cost ratio for a binomial sequential sampling plan is calculated by the following equation:

$$\text{benefit-cost ratio} = [\Sigma P_c / (n * c)] * 100$$

where P_c is the sum of proportional correct decisions, n is the average number of samples required to make a decision, and c is the total cost of collecting the sample, as in relative net precision. In addition to the estimating sampling cost, the frequency (%) of correct treat and no-treat decisions also should be taken into consideration when cost values are similar (Burkness et al. 1999). If sampling plans have similar frequencies of correct decisions, but one has lower cost, the benefit-cost ratio will be greater for the plan with reduced cost. If two plans have similar ratios, the OC values at the threshold should be compared. Ideally, the OC value at the threshold should be 0.50, signifying an equal probability of treating or not treating at the action threshold. For example, if the probability of treating at the action threshold was 0.40, the sampling plan would recommend a treatment too often (i.e., a conservative plan).

Results and Discussion

Enumerative Sampling. A regression analysis of the log-mean and log-variance showed a positive linear correlation with a slope of 1.54 ± 0.04 (SE), an intercept of 0.96 ± 0.06 (SE), $R^2 = 0.94$, Taylor's $a = 9.16$, and Taylor's $b = 1.54$. The b value was significantly $>$ 1 ($t = 12.33$, df = 87, $P < 0.0001$), suggesting a highly aggregated spatial distribution for *A. glycines.* Initial results from Green's plan indicated the mean actual precision levels were higher than desired at both precisions of 0.10 and 0.25. Initial resampling analysis for the desired precision of 0.10, without replacement, resulted in an actual average precision of 0.13 and an actual average sample number of 134 individual plants per field (data not shown). Initial resampling analysis for the desired precision of 0.25, without replacement, resulted in an actual average precision of 0.27 and an actual average sample number of 33 plants per field (data not shown).

Sample size requirements for the desired precision level of 0.10 and 0.25 declined rapidly as mean density increased (Fig. 2). The maximum mean sample size for densities of three aphids per plant at a precision of 0.10 was ≈870 (Fig. 2A), and the maximum mean sample size for densities of aphids at a precision of 0.25 was ≈105 (Fig. 2B). As the mean density of *A. glycines* increased to ≈125, mean sample size decreased at both 0.10 (130 samples) and 0.25 (15 samples) precision levels.

Hutchison et al. (1988) and Naranjo and Flint (1995) reported that the fixed precision level could be relaxed (i.e., decreasing the initial desired precision) to further reduce the number of samples required for sampling while maintaining the desired precision. We subsequently refined our fixed precision levels to attain actual 0.10 and 0.25 precision levels by increasing the desired precision. Our final adjustment, using Resampling for Validation of Sample Plans, was to increase precision levels to 0.08 and 0.23 for ecological and pest management purposes, respectively. After this modification, final validation resulted in acceptable average precision for both types of research (Tables 1 and 2). Validation of the 15 data sets without replacement at 0.10 was not possible because some data sets had very low densities. The resampling analysis was reformatted to run with replacement where all data sets were successfully processed (Naranjo and Hutchison 1997). Final simulations, to achieve actual precision levels of 0.10 and 0.25, required average sample numbers of 310 and 38 individual plants, respectively (Tables 1 and 2).

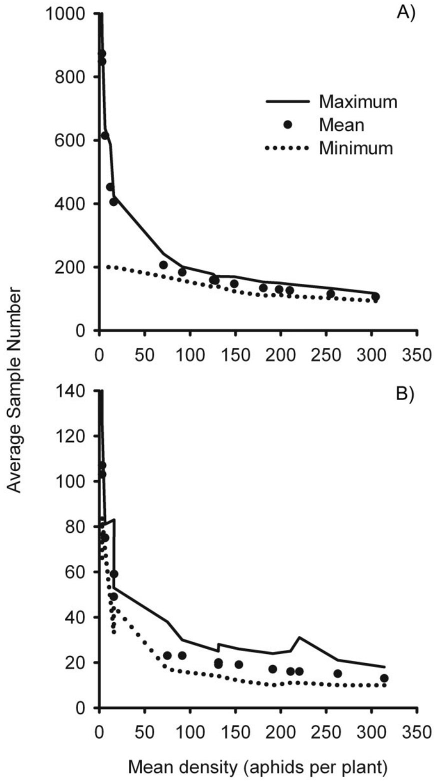

Fig. 2. Summary of resampling analysis showing actual average sample number for Green's sequential sampling plan for a range of *A. glycines* densities. (A) Green's plan with preset precision, 0.08 (actual mean precision, 0.10). (B) Green's plan with preset precision, 0.23 (actual mean precision, 0.25).

Binomial Sampling. The OC function curves for all calculated plans were plotted (Fig. 3A–J). At the optimal tally threshold and action threshold combination (tally threshold, 40; mean density, 250 per plant; and action threshold, 0.84), the OC was 0.50 (Fig. 3J). This result indicates this sampling plan is neither conservative nor liberal. A less than optimal combination (tally threshold, 20; mean density, 250 per plant; action threshold, 0.89), resulted in a liberal OC of 0.60 (Fig. 3E), where the plan is more likely to recommended not to treat than to treat at the threshold.

Resampling analysis for Wald's binomial sampling plan resulted in 10 summary outputs. These results were used to create the treatment probability matrix for determining the optimal tally threshold and action threshold combination. The overall proportion of correct treat and no-treat decisions (i.e., $A_i + D_i$ values) was apparent for all combinations (Table 3). In addition and perhaps more importantly, the matrix also summarizes the proportion of incorrect decisions (i.e., $B_i + C_i$) for each combination. The tally threshold and action threshold of 40 and 0.837 (84% of plants infested with ≥40 aphids each), respectively, proved the most effective sampling plan in reducing incorrect no-treat decisions (Table 3), and is consistent with the currently accepted economic threshold of 250 aphids per plant. The corresponding minimum number of samples for the optimal tally threshold and action threshold combination was 11 plants (Table 3).

Final stop lines for the binomial sampling plan, with α and β values <0.10, can be used in the field (Fig. 4). Initially, 11 plants are sampled and each plant is determined as not infested (<40 aphids) or infested (≥40 aphids). The percentage of infested plants is then calculated from the 11-plant sample and compared with the stop line graph (Fig. 4). An immediate management decision can be made if the percentage is in the "treat" or "do not treat" area. If the value is between the stop lines, additional samples will be required before making a treatment decision. We recommend not taking >31 plant samples to reach a decision for the binomial sequential plan and to resample that same field in 3–4 d if the estimate is still within the no-decision zone. The suggested resample period is based upon the knowledge of the reproductive capacity of the *A. glycines* and the potential to double populations in 1.5 d in laboratory conditions (McCornack et al. 2004).

Chinese literature has demonstrated sampling techniques for *A. glycines,* but these plans may not be appropriate for North American soybean. Liu (1986) and Su et al. (1996) both suggest using sequential sampling plans to estimate *A. glycines* densities in China. Both also describe *A. glycines* as an aggregated species, but analyses were not directly comparable to the Taylor's power law *b* value. However, economic thresholds were never defined and sampling plans were not validated for accuracy. Huang et al. (1992) discuss various sampling techniques, but do not provide the average sample number or a degree of precision, and do not validate data sets for a range of density estimates.

Cost Analysis. The enumerative plan was derived from whole plant counts of *A. glycines,* and therefore estimates of cost were based on sampling whole plants. From data sets collected by experienced samplers in 2001–2003, plant sampling time ranged from 0.25 to 2 min per plant, for densities of 0–2,700 aphids per plant, respectively. The 3-yr season average for sampling was 60 s per plant with an average of 30 s walking between plants (E.W.H., unpublished data). With the known minimum sample required for the enumerative and binomial sequential sampling plans, cost efficiency could then be estimated. The relative net precision for the fixed precision sampling plans was highest for IPM (0.25) with an average sampling time of 57 min per field, compared with the relative net precision for research (0.10) averaging 465 min per field (Table 4). Even though a research-based enumerative plan provides density estimates closer to the actual mean density, the cost of obtaining the improved estimate could be prohibitive for some pest management applications.

Table 1. Resampling simulations used to validate fixed precision sampling plan [Green (1970)] for *A. glycines* by using a preset precision level of 0.08 (desired = 0.10), with replacement (Taylor's $a = 9.157$, and $b = 1.543$)

Validation data set	Observed mean density	Avg statistics for 500 sequential sampling simulations[a]						
		Density	Precision			Avg sample no.		
		Mean	Mean	Min.	Max	Mean	Min.	Max
1	3.10	3.00	0.11	0.06	0.16	873	200	1017
2	3.15	3.15	0.06	0.06	0.07	848	200	934
3	6.35	6.38	0.03	0.02	0.03	614	200	637
4	12.82	13.21	0.24	0.04	0.28	452	200	587
5	15.95	15.90	0.03	0.03	0.03	405	200	424
6	70.67	70.77	0.12	0.09	0.15	206	170	242
7	90.21	91.42	0.09	0.06	0.12	183	158	201
8	125.13	125.54	0.11	0.09	0.13	159	137	178
9	127.73	127.97	0.07	0.06	0.09	157	141	171
10	147.39	148.85	0.13	0.08	0.16	147	123	170
11	181.90	180.75	0.12	0.08	0.15	134	110	153
12	196.51	198.22	0.13	0.09	0.15	129	112	150
13	208.48	209.89	0.12	0.10	0.14	126	108	146
14	253.63	255.16	0.10	0.09	0.12	115	102	133
15	303.46	304.62	0.08	0.06	0.10	106	93	117
Overall	116.43	116.99	0.10	0.07	0.13	310	150	351

[a] Data set resampled with replacement because of low mean *A. glycines* densities in Resampling for Validation of Sample Plans simulation software (Naranjo and Hutchison 1997).

Overall, the binomial sequential sampling plans were more efficient. A management decision could potentially be made in <10 min by using tally thresholds of 20 aphids or 40 aphids to estimate *A. glycines* infestations (Table 4). Although the tally threshold of 20 aphids requires less sampling time and has a higher benefit-cost ratio, we recommend the tally threshold of 40 aphids because of a higher frequency of correct treat and no-treat decisions and the OC value at the threshold was more conservative (Table 3; Fig. 1). Sampling soybean plants with tally thresholds of ≥20 and ≥40 aphid averages 10.5 s per plant with 30-s walking time between plants, and 17.3 s per plant with 30-s walking time between plants, respectively (E.W.H., unpublished data). Average time estimates were collected over a range of experienced and novice *A. glycines* samplers, and is considered applicable to all potential users.

The purpose of this article was to develop a preliminary sampling protocol for a new insect pest of commercial soybean. We also wanted to incorporate multiple data sets that included a wide range of *A. glycines* densities. Although an economic threshold for *A. glycines* has only been tentatively established, designing a plan to estimate densities is essential to improve our understanding of density/yield relationships and to provide guidelines for IPM.

We developed an enumerative sampling plan for estimating *A. glycines* densities for two levels of precision. Taylor's b value indicated the *A. glycines* was

Table 2. Resampling simulations used to validate fixed precision sampling plan [Green (1970)] for *A. glycines* by using a preset precision level of 0.23 (desired, 0.25) without replacement (Taylor's $a = 9.157$ and $b = 1.543$)

Validation data set	Observed mean density	Avg statistics for 500 sequential sampling simulations[a]						
		Density	Precision			Avg sample no.		
		Mean	Mean	Min.	Max	Mean	Min.	Max
1	3.10	3.24	0.28	0.14	0.57	107	64	141
2	3.15	3.24	0.18	0.14	0.24	103	84	126
3	6.35	6.43	0.07	0.06	0.09	75	70	81
4	12.82	16.29	0.37	0.10	0.84	59	33	83
5	15.95	16.03	0.09	0.07	0.11	49	45	53
6	70.67	75.00	0.33	0.17	0.56	26	17	38
7	90.21	91.43	0.23	0.08	0.45	23	16	30
8	125.13	131.45	0.30	0.17	0.46	20	14	28
9	127.73	131.26	0.20	0.08	0.30	19	14	25
10	147.39	153.85	0.30	0.13	0.63	19	12	26
11	181.90	191.18	0.30	0.18	0.62	17	10	24
12	196.51	210.78	0.34	0.16	0.54	16	11	25
13	208.48	220.56	0.32	0.16	0.56	16	11	31
14	253.63	262.91	0.28	0.14	0.43	15	10	21
15	303.46	314.08	0.21	0.08	0.37	13	10	18
Overal	116.43	121.85	0.25	0.12	0.45	38	28	50

[a] Data set resampled without replacement in Resampling for Validation of Sample Plans simulation software (Naranjo and Hutchison 1997).

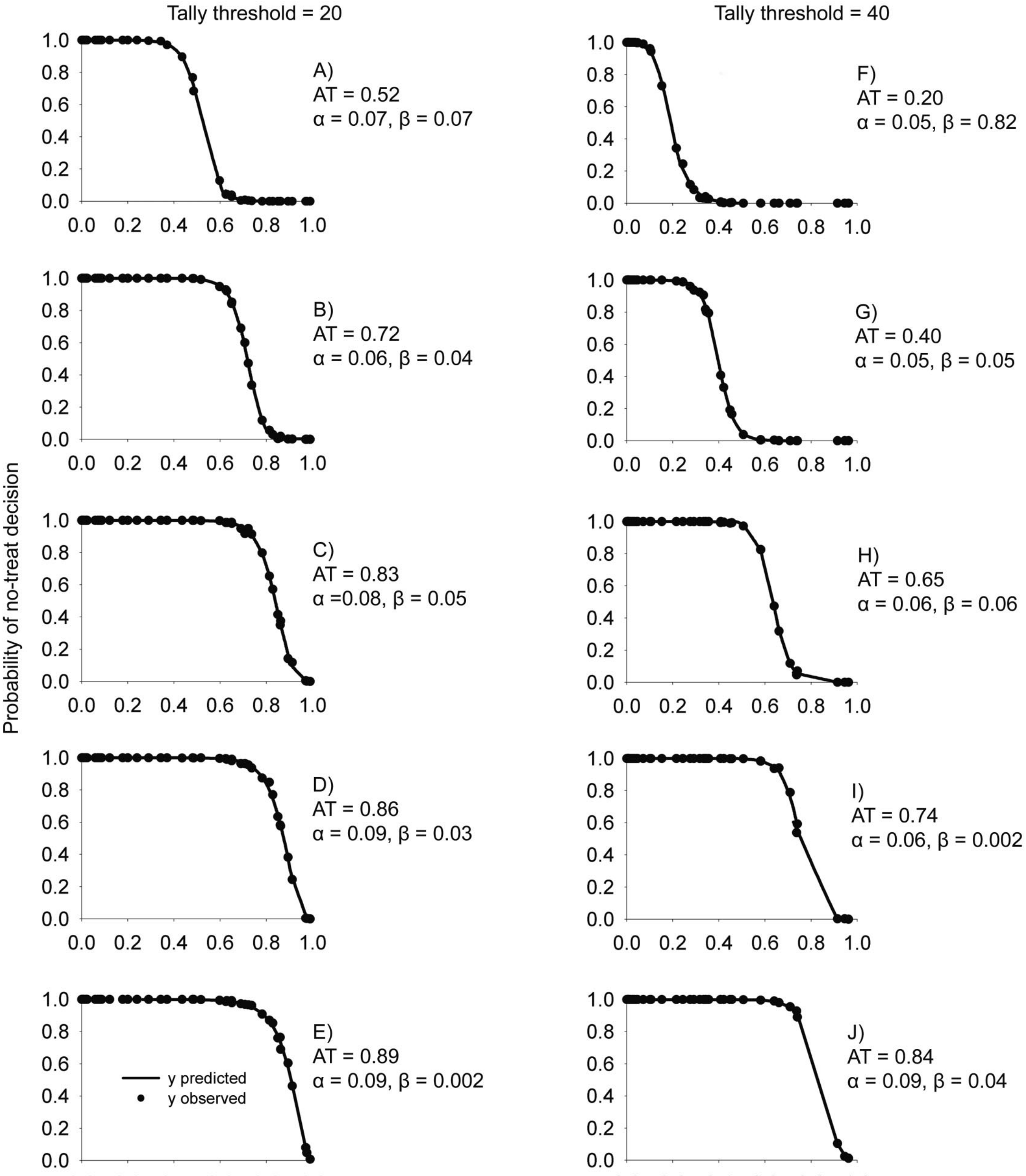

Fig. 3. OC function for binomial sequential sampling plans for *A. glycines*. The predicted line (solid line) was fitted using $y^{-1} = a + bx + c/x^{1.5}$ for all graphs. The OC was plotted against the observed proportion infested obtained from resampling software; AT, action threshold and actual α and β values are shown, based on preset values of $\alpha = 0.10$ and $\beta = 0.10$ (also see Table 3).

spatially aggregated, and this behavior may be especially apparent at low densities. As a result, taking >300 samples is not a practical, but intensive sampling (≈7.75 person-hours) may be required for estimating populations of *A. glycines* in ecological studies when densities are low (Tables 1 and 4). However, our results for estimating densities for IPM research are realistic (≈1 person-hour), and averaged just 38 plants over a wide range of *A. glycines* densities (Tables 1 and 4).

Table 3. Probability of correct and incorrect treatment decisions by using selected tally thresholds for a binominal sequential sampling plan for *A. glycines* based on the mean, the estimated proportion of plants infested, and operating characteristic

$\bar{x}$ Aphids[a]	Action threshold[b]	α, β[c]	Correct decisions		Incorrect decisions		Avg sample no.
			A[d]	D[e]	B[f]	C[g]	
			Tally threshold, 20				
25	0.516	0.069, 0.072	0.283	0.703	0.005	0.010	12.58
50	0.715	064, 0.040	0.161	0.805	0.014	0.018	13.11
100	0.830	0.080, 0.048	0.092	0.873	0.019	0.018	10.99
150	0.864	0.088, 0.033	0.060	0.892	0.022	0.030	11.01
250	0.889	0.088, 0.002	0.051	0.912	0.016	0.020	10.41
			Tally threshold, 40				
25	0.196	0.046, 0.820	0.311	0.670	0.013	0.005	11.49
50	0.400	0.053, 0.052	0.190	0.783	0.016	0.012	13.49
100	0.636	0.056, 0.059	0.094	0.886	0.014	0.005	12.45
150	0.744	0.055, 0.002	0.041	0.942	0.000	0.017	11.30
250	0.837	0.088, 0.044	0.039	0.955	0.002	0.004	10.33

[a] Action threshold (proportion of plant infested derived from mean aphid densities) values calculated from the proportion of plants infested and mean relationship curves shown in Fig. 3.

[b] Five mean values chosen to represent a range of possible *A. glycines* economic threshold densities.

[c] The α error rate (type I) describes the probability of treating a field when the actual pest density is below the action threshold, and the β error rate (type II) describes the probability of not treating a field when the actual pest density exceeds the action threshold. All α and β values were preset at 0.10 for resampling simulations; final actual α and β values were estimated from the fitted curves in Fig. 3.

[d] A is the probability of both the estimated proportion infested and the true pop being above their respective action thresholds leading to a correct treat decision.

[e] D is the probability of both the estimated proportion infested and the true pop being below their respective action thresholds, leading to a correct no-treat decision.

[f] B is the probability of the estimated proportion infested being below the proportion infested action threshold, and the true pop is above the density action threshold, leading to an incorrect no-treat decision.

[g] C is the probability of the estimated proportion infested being above the proportion infested action threshold, and the true pop is below the density action threshold, leading to an incorrect treat decision.

In addition to the enumerative plan, the binomial sampling protocol can be used for improved efficiency in IPM decision making. With this plan, a minimum of only 11 plant samples are required with average sample numbers ranging from 11 to 19 total plants over the density range we studied. As with all sequential plans, the highest numbers of samples required to make a treatment decision occur near the economic threshold. The plan will provide the greatest saving when *A. glycines* infestation levels are above or below the economic threshold. The proposed binomial sampling plan (tally threshold, ≥40; action threshold, 0.84) reflects a current economic threshold of 250 aphids per plant (Table 2). This plan does not take into consideration higher economic thresholds that may be developed for soybean plants at a later stage of maturity (R5 and later), resistant varieties, or later planting dates, where yield potential may be reduced. The binomial sampling plan, however, should help crop managers make treatment decisions quickly, before aphids reach economically damaging levels for most Midwestern soybean production systems.

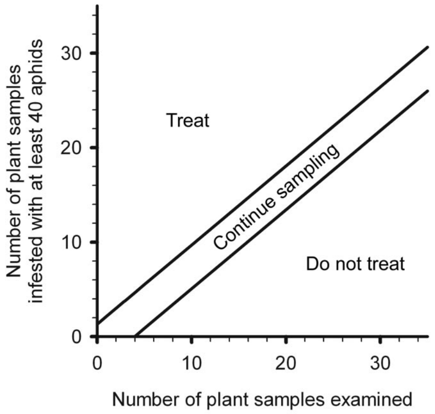

Fig. 4. Decision stop lines for the binomial sequential sampling plan for *Aphis glycines* based on resampling analysis, where actual $\alpha = 0.055$, $\beta = 0.002$, a tally threshold of 40 (≥40 aphids per plant to be considered infested), and an action threshold of 0.84.

The enumerative and binomial sequential sampling plans for *A. glycines* were created from data sets collected in Minnesota; however, the potential applications to other soybean growing regions is unknown. Possible refinements of each sampling plan could include using a smaller sample unit (i.e., leaflet or trifoliate instead of whole plant counts), or developing a dynamic sampling plan as the plant matures.

Acknowledgments

We thank Anne-Marie Hanson, Brian McCornack, and Adriano Pereira for assistance in the field, and Steve Naranjo (Western Cotton Research Laboratory, UDSA–ARS, Phoenix, AZ) for software consultation. We also are grateful to Robert Venette and Robert Koch (Department of Entomology, University of Minnesota) for reviewing previous drafts of this manuscript. This work was funded by the North Central Soybean Research Program, the Minnesota Soybean Research and Promotion Council, and the University of Min-

Table 4. Comparing sampling efficiency cost (in time) using four plans created with Green's enumerative fixed-precision and Wald's binomial sequential sampling plans for *Aphis glycines*

Sampling plan	Minimum sample no.	Avg sample time per plant (min)	Total sample time (min)[a]	RNP[b]	BCR[c]
Enumerative[d]					
Fixed precision, 0.10	310	1.00	465	1.29	
Fixed precision, 0.25	38	1.00	57	4.21	
Binomial sequential[e]					
TT, 20; AT, 0.889[f]	11	0.17	7.4		13.07
TT, 40; AT, 0.837[g]	11	0.28	8.6		11.58

[a] Time calculated by [min. sample no. x (avg sample time per plant in minutes + 0.5-min walking between plant samples)].

[b] Relative net precision = $[(1/(RV * c)) * 100]$, where RV is $(se/\bar{x}) * 100$, and c is the total cost related to collecting the given number of samples usually measured in person-hours; relative net precision calculations are not possible with binomial plans because precision is not defined.

[c] Benefit-cost ratio = $[(\Sigma P_c / (n * c)) * 100]$, where P_c is the sum of proportional correct decisions, n is the average number of samples required to make a decision, and c is the total cost of collecting the sample; benefit-cost ratio calculations are not possible with enumerative plans because percentage of correct decisions is not defined.

[d] Enumerative plan was based on whole plant counts of *A. glycines.*

[e] Binomial sequential plans are based on tally thresholds, or a minimum number of aphids on a plant to be considered infested.

[f] TT, tally threshold; AT, action threshold; binomial sequential plan based on a tally threshold of ≥20 *A. glycines* to consider a plant infested, and an action threshold of 89% (=250 aphids per plant).

[g] TT, tally threshold; AT, action threshold; binomial sequential plan based on a tally threshold of ≥40 *A. glycines* to consider a plant infested, and an action threshold of 84% (=250 aphids per plant).

nesota College of Agriculture, Food and Environmental Sciences Rapid Agricultural Response Fund.

References Cited

Burkness, E. C., and W. D. Hutchison. 1997. Development and validation of a binomial sequential sampling plan for striped cucumber beetle (Coleoptera: Chrysomelidae) in cucurbits. J. Econ. Entomol. 90: 1590–1594.

Burkness, E. C., and W. D. Hutchison. 1998. Development and validation of a fixed-precision sampling plan for estimating striped cucumber beetle (Coleoptera: Chrysomelidae) density in cucurbits. Environ. Entomol. 27: 178–183.

Burkness, E. C., R. C. Venette, P. K. O'Rourke, and W. D. Hutchison. 1999. Binomial sequential sampling for management of aster leafhopper (Homoptera: Cicadellidae) and aster yellows phytoplasma in carrot: impact of tally threshold on the accuracy of treatment decisions. Environ. Entomol. 28: 851–857.

Calvin, D. D., M. C. Knapp, K. Xingquan, F. L. Poston, and S. M. Welch. 1986. Using a decision model to optimize European corn borer (Lepidoptera: Pyralidae) egg-mass sampling. Environ. Entomol. 15: 1212–1219.

Elliott, N. C., K. L. Giles, T. A. Royer, S. D. Kindler, F. L. Tao, D. B. Jones, and G. W. Cuperus. 2003. Fixed precision sequential sampling plans for the greenbug and bird cherry-oat aphid (Homoptera: Aphididae) in winter wheat. J. Econ. Entomol. 96: 1585–1593.

Giles, K. L., T. A. Royer, N. C. Elliott, and S. D. Kindler. 2000. Development and validation of a binomial sequential sampling plan for the greenbug (Homoptera: Aphididae) infesting winter wheat in the southern plains. J. Econ. Entomol. 93: 1522–1530.

Green, R. H. 1970. On fixed precision sequential sampling. Res. Popul. Ecol. 12: 249–251.

Hamilton, G. C., J. H. Lashomb, S. Arpaia, R. Chianese, and M. Mayer. 1998. Sequential sampling plans for Colorado potato beetle (Coleoptera: Chrysomelidae) in eggplant. Environ. Entomol. 27: 33–38.

Hamilton, A. J., N. A. Schellhorn, N. M. Endersby, P. M. Ridland, and S. A. Ward. 2004. A dynamic binomial sequential sampling plan for *Plutella xylostella* (Lepidoptera: Plutellidae) on broccoli and cauliflower in Australia. J. Econ. Entomol. 97: 127–135.

Hill, J. H., R. Alleman, D. B. Hogg, and C. R. Grau. 2001. First report of transmission of soybean mosaic virus and alfalfa mosaic virus by *Aphis glycines* in the New World. Plant Dis. 85: 561.

Huang, F., X. Ding, X. Wang, and Z. Huang. 1992. Studies on the spatial distribution pattern of soybean aphid and sampling techniques. J. Shenyang Agric. Univ. 23: 81–87.

Hutchison, W. D. 1994. Sequential sampling to determine population density, pp. 207–244. *In* L. Pedigo and G. Buntin [eds.], Handbook of sampling methods for arthropods in agriculture. CRC, Boca Raton, FL.

Hutchison, W. D., D. B. Hogg, M. A. Poswal, R. C. Berberet, and G. W. Cuperus. 1988. Implications of the stochastic nature of Kuno's and Green's fixed-precision stop lines: sampling plans for the pea aphid (Homoptera: Aphididae) in alfalfa as an example. J. Econ. Entomol. 81: 749–758.

Jones, V. P. 1994. Sequential estimation and classification procedures for binomial counts, pp. 175–205. *In* L. Pedigo and G. Buntin [eds.], Handbook of sampling methods for arthropods in agriculture. CRC, Boca Raton, FL.

Landis, D., M. Brewer, and G. Heimpel. 2003. Soybean Aphid and Parasitoids 2003: a survey of NCR-125 cooperating states. *In* 2003 Michigan state report to NCR-125. http://www.cips.msu.edu/ncr125/StateRpts2003MI.htm (9 April 2004).

Liu, Z. 1986. Spatial distribution pattern and sequential sampling of soybean aphid. Plant Protection 12: 16–18.

McCornack, B. P., D. W. Ragsdale, and R. C. Venette. 2004. Demography of soybean aphid (Homoptera: Aphididae) at summer temperatures. J. Econ. Entomol. 97: 854–861.

[MPR] Minnesota Pest Report. 2003. Soybean aphid. MPR 11–03. http://www.mda.state.mn.us/pestsurvey/PestReports/2003aug08.pdf (15 May 2004).

Naranjo, S. E., and H. M. Flint. 1995. Spatial distribution of adult *Bemisia tabaci* (Homoptera: Aleyrodidae) in cotton and development and validation of fixed-precision sampling plans for estimating population density. Environ. Entomol. 24: 261–270.

Naranjo, S. E., and W. D. Hutchison. 1997. Validation of arthropod sampling plans using a resampling approach: software and analysis. Am. Entomol. 43: 48–57.

[NCSRP] North Central Soybean Research Program. 2004. Soybean aphid: research update. Plant Health Initiative. http://www.planthealth.info/aphidinsert/aphidinsert.htm (6 May 2004).

Nyrop, J. P., and M. Binns. 1991. Quantitative methods for designing and analyzing sampling programs for use in pest management, pp. 67–132. *In* D. Pimentel (ed.), Handbook of pest management in agriculture. CRC, Boca Raton, FL.

Onsager, J. A. 1976. The rationale of sequential sampling, with emphasis on its use in pest management. U.S. Dep. Agric. Tech. Bull. 1526.

Ostlie, K. 2001. Soybean aphid reduces yields: harvest results from insecticide strip trials. University of Minnesota, St. Paul, MN. http://www.soybeans.umn.edu/crop/insects/aphid/ studyresults/htm (27 January 2004).

Pedigo, L. P., G. L. Lentz, J. D. Stone, and D. F. Cox. 1972. Green cloverworm populations in Iowa soybean with special reference to sampling procedure. J. Econ. Entomol. 65: 414–421.

Ragsdale, D. W., D. J. Voegtlin, and R. J. O'Neil. 2004. Soybean aphid biology in North America. Ann. Entomol. Soc. Am. 97: 204–208.

Silva, L., V. Virgilio, and J. Tavares. 2003. Sampling plans for *Pseudaletia unipuncta* (Lepidoptera: Noctuidae) larvae in Azorean pastures. Environ. Entomol. 32: 1211–1218.

Southwood, T.R.E. 1978. Ecological methods, with particular reference to the study of insect populations. Chapman & Hall, London, United Kingdom.

Su, J., K. Hai, and X. Shi. 1996. Spatial distribution and sampling techniques of *Aphis glycines* Matsumura. J. Nanjing Agric. Univ. 19: 55–58.

Taylor, L. R. 1961. Aggregation, variance and the mean. Nature (Lond.) 189: 732–735.

Venette, R. C., and D. W. Ragsdale. 2004. Assessing the invasion by soybean aphid (Homoptera: Aphididae): where will it end? Ann. Entomol. Soc. Am. 97: 219–226.

Wald, A. 1947. Sequential analysis. Wiley, New York.

Ward, S. A., R. Rabbinge, and W. P. Mantel. 1985. The use of incidence counts for estimation of aphid populations. 2. Confidence intervals from fixed sample sizes. Neth. J. Plant Pathol. 91: 100–104.

Ward, S. A., R. J. Chambers, K. Sunderland, and A.F.G. Dixon. 1986a. Cereal aphid populations and the relation between mean density and spatial variance. Neth. J. Plant Pathol. 92: 127–132.

Ward, S. A., K. D. Sunderland, R. J. Chambers, and A.F.G. Dixon. 1986b. The use of incidence counts for estimation of cereal aphid populations. 3. Population development and the incidence-density relation. Neth. J. Plant Pathol. 92: 175–183.

Wilson, L. T., C. Pickel, R. C. Mount, and F. G. Zalom. 1983. Presence-absence sequential sampling for cabbage aphid and green peach aphid (Homoptera: Aphididae) on brussel sprouts. J. Econ. Entomol. 76: 476–479.

Wu, Z., D. Schenk-Hamlin, W. Zhan, D. W. Ragsdale, and G. E. Heimpel. 2004. The soybean aphid in China: a historical review. Ann. Entomol. Soc. Am. 97: 209–218.

Received 15 May 2004; accepted 21 July 2004.