Sampling Techniques for the Evaluation of *Aphis gossypii* Glover (Homoptera: Aphididae) Infestation in Cotton, *Gossypium hirsutum* (L.)

S.M. Mazza[1], G.B. Contreras[2], M.A. Simonella[2], M.G.A. Polak[2], J.A. Schroeder[1], C.J. Tannure[1] and O.M. Royo[2]

[1] UNNE – Corrientes, Argentina
[2] INTA - Saenz Peña, Chaco, Argentina

ABSTRACT

In the search for a sampling technique to get better accuracy and lower cost in assessing aphid infestation in cotton crops, weekly systematic random sampling was undertaken in two 50 ha fields in 1995/96. The number of aphids per plant, the proportion of plants attacked during the crop cycle and number of aphids in the top six nodes between the seventh and ninth week after planting were recorded. The following evaluations were undertaken: 1) comparison of the efficacy of two sampling methods was examined: a) random systematic sampling with one plant per sampling point (290 plants); and b) cluster system with 10 plants at each sampling point (29 clusters); 2) the association between number of aphids per plant and proportion of attacked plants; 3) the association between the number of aphids in the top six nodes and those in the whole plant. In both fields, the cluster method of sampling was more efficient (higher precision with less cost) than the single unit method. The number of aphids per plant and proportion of attacked plants were not significantly correlated (p > 0.05), indicating the risk of taking decisions based on one without knowing the other. The number of aphids in the top six nodes and in the whole plant is significantly different but significantly correlated. The top of the plant is more strongly infested but the total infestation could be estimated from the infestation in the top six nodes.

Introduction

In order to make appropriate pest control decisions in cotton crops, periodic evaluations of presence of pests and level of infestation are necessary (Barral and Zago, 1983). Complete insect population census is practically impossible so, infestations must be estimated from samples to establish certain parameters (means, variances, totals, proportions) with an accuracy that can be determined by the bound of the error of estimation (Cochran, 1974; Cochran, 1983; Schaeffer *et al.*, 1979). Sampling should aim to meet the information needs of the researcher with minimum effort. Variation in the data can be controlled by designing the sample survey according to sampling theory (Cochran, 1974; Schaeffer *et al.*, 1979).

Sampling estimators are random variable functions; their probabilistic distributions depend on the sampling design and the population and sampling sizes (Schaeffer *et al.*, 1979; Southwood, 1995). The function of a parameter estimator is to generate values concentrated on the parameter and preferably close to it. A sampling programme that secures an estimator with a mathematical expectancy equal to the parameter and with low variability has to be selected (Cochran, 1974; Schaeffer *et al.*, 1979; Steel and Torrie, 1988). If the aim is to obtain estimates of the mean density, then it is desirable to minimize the variance, but if the pattern of dispersion of animals is the principal interest, then a small variance is not necessary (Southwood, 1995).

Cluster sampling gives more information per unit cost than simple or stratified random sampling when the cost of obtaining observations increases as the distance separating the elements to be sampled increases and gives more precision when the variability within clusters is greater than between clusters (Cochran, 1974; Schaeffer *et al.*, 1979; Steel and Torrie, 1988).

Population dispersion has a considerable significance because it not only affects the sampling programmes and the methods of data analysis, but can be used to give a measure of population size. Understanding dispersion is also vital in the analysis of predator-prey and host-parasite relationships (Southwood, 1995). In ecological studies, the variance is commonly found to be larger than the mean, that is the distribution is contiguous and the population is clumped or aggregated (Ibid.). Changes in the population density often signify changes in the distribution; usually random patterns are expected when densities are low and contiguous patterns when the density increases (Pielou, 1977; Rogers, 1974; Southwood, 1995).

It is often expensive to estimate the actual number of insects in a sample, but their presence or absence can be easily assessed. If the dispersion of such an animal in a particular habitat can be described by negative binomial distribution, the probability of a particular mean population can be estimated from the presence - absence proportion (p). This is particularly useful with highly clumped animals. The saving on counting time is important. The sensitivity is also important and is reliable only if the critical density levels are related to values of p less than about 0.8. Above this the uncertainty associated with predictions is too great (Southwood, 1995).

Aphid populations over plants are not randomly distributed in relation to height. At the beginning and end of the season, the upper part of the plant is more heavily infested than the lower. This relationship changes in mid-summer (Kapatos *et al.*, 1996). The aphid infestation level in cotton crops can be estimated with accuracy through field examination (Williams *et al.*, 1995). Elberson and Johnson (1995) compared field examination with Berlese's extraction and the suction trap sample. Hardee *et al.* (1994) suggested the use of whole plant washing in the first three weeks after planting, sampling the fourth fully expanded leaf from the terminal from the fourth through the sixth week after planting and sampling the first main stem green leaf about one-third the distance from the terminal from the seventh through the ninth week after planting. The rest of the season, sample the first main stem green leaf above the first basal fruiting branch.

With the objective of carrying out insect control in commercial plantations, periodic evaluations of presence and pest levels are necessary (Barral and Zago, 1983). In northern plantations in Argentina, evaluations are generally made by field examination by non-random sampling, selecting plants systematically in a transect line or looking for plants near the lines head. The variables determined in most cases are the presence or absence of insects and the proportion of infested plants. Because of the high cost of counting insects, control decisions are based on the infestation level estimated by these methods. These decisions can be ill judged if made on the basis of a non random selection of plants, or ignoring the means distributions and the bound of the error of estimation because there may be an important bias in the estimates (Cochran, 1983).

Based on the methods of evaluation in use in the region and looking for practical solutions applicable to extensive plantations, our objective is to establish a sampling methodology that allows estimation of the infestation level with more precision than now; evaluating this precision by the bound of the error of estimation. We worked with cotton aphid infestation in two seeding systems (direct cotton seeding over wheat and conventional). Some aspects of sampling were contemplated: 1) The sampling method according to the hypothesis that the cluster is a more efficient sample than simple units if the cost per unit increases with the distance and most accurate if the variability is higher within than between clusters; 2) The parameters to be estimated comparing the proportion of infected plants with the mean of aphids by plant, under the hypothesis that when two measures are significantly correlated, one of them can be estimated by the other; and 3) The possibility of counting only the upper six nodes and then estimating the whole plant infestation, under the same hypothesis as in 2.

Materials and methods

Field work was done in two 50 ha, furrow irrigated, cotton fields with direct seeding over wheat (A) and conventional seeding (B), in Salta (Argentina) during 1995/96. The sampling population was all the plants in the 50 ha unit, approximately 5,000,000. Systematic random samplings were designed: with one plant per sampling point (S) (290 plants) and by clusters of 10 plants at each sampling point (C) (29 clusters) (Cochran, 1974; Cochran, 1983; Schaeffer *et al.*, 1979). With both systems, the number of aphids in the whole plant (NP) and the proportion of plants infested by aphids (PPI) were evaluated weekly during the entire crop cycle. In addition, the number of aphids at the top six nodes (NU) were evaluated from the seventh to the ninth week after planting.

The population estimators for the numbers of aphids per plant and for variability were determined and the bound of the error of estimation at p = 0.05 were calculated for both sampling designs (Schaeffer *et al.*, 1979). A comparative analysis of variability and accuracy between S and C designs was carried out (Cochran, 1974; Schaeffer *et al.*, 1979; Steel and Torrie, 1988). In order to determine the aphids' spatial distribution pattern, the relation between mean and standard deviation was analyzed (Southwood, 1995).

The aphid population distribution over the plant was analyzed comparing NU and NP using the "t" test (Steel and Torrie, 1988). Pearson correlation coefficient (r) was calculated evaluating the possibility of estimating the NP from NU and NP by PPI, the latter only for the single points sampling design with 290 plants (Steel and Torrie, 1988).

Results and discussion

The number of aphids per plant (NP) in both lots, shows a considerable variation over the crop cycle (the mean (M) in seeding over wheat from 0.114 to 181.297 and in direct seeding from 0.669 to 317.454) being higher in the intermediate cotton cycle states in both crop systems. The standard deviation (SD) has a similar variation, giving the relationship SD/M > 1 for single units sampling. This indicates a contiguous spatial distribution, denoting an aggregate population according to Southwood (1995). The bound of the error of estimation oscillated between 3.01 and 2,708 for S, between 0.096 and 138 for C (lot A), between 15.95 and 5,834for S, and between 0.136 and 161.8 for C (lot B) (Table 1 and Figures 1 and 2), denoting that for estimating number of aphids per plant, in both crop systems, the cluster sampling method gives more accuracy.

The same variation is seen in the proportion of infested fields as in the total number of aphids (M between 0.069 and 0.983, for lot A, and between 0.123 and 0.954 for lot B). For this variable the cluster design shows lower SD and BEE (lot A: between 0.0006 and 0.01; lot B: between 0.00007 and 0.0385) than the single units one (Table 1 and Figures 3 and 4).

In both seeding systems (direct over wheat and conventional) and variables analyzed (number of aphids by plant and proportion of infested plants) the cluster sampling design with 10 plants by cluster had brought lower bound of the error of estimation than the single units design. These results agree with Scheaffer *et al.* (1979) and Steel and Torrie (1988) because variation within clusters is larger than between. In this case, cluster sampling is more efficient than the single units because it can control the variation between sampling units and each cluster exhibits the population variability.

The "t" test between NU and NP was significant (lot A $t = 2.20$ and lot B $t = 2.32$; $\alpha = 0.05$), with the conclusion that the upper part of the plant is more heavily infested than the lower, according to Kapatos *et al.* (1996). The Pearson correlation coefficient was 0.846 and significant ($p = 0.05$), permitting estimates of the whole plant infestation from the upper six nodes, according with Hardee *et al.* (1994).

Pearson correlation coefficients (r) between the number of aphids per plant and the proportion of plants infested were 0.50 in lot A and 0.22 in lot B, both not significant (Figures 5 and 6). It can be observed generally that both variables increase simultaneously but not linearly, necessitating the study of nonlinear functions. In some situations, the proportion of infested plants can give low values with large number of aphids per plant; leading to underestimated of the infestation, not only when the proportion of infested plants is above 0.8 but also when it is less, as Southwood (1995) pointed out. This indicates the risk of control plans using one of the variables without knowing the other. The variable must be selected according to the relative importance of determining the real number of aphids or only evaluating the proportion of infestation and their effects on the crop.

Some recommendations can be made n the selection of appropriate sampling designs for the evaluation of aphid infestation in cotton crops:

1- Use the cluster sampling method because the aphid spatial dispersion generates high variability between plants that are close together.

2- Evaluate the number of aphids per plant but do not estimate it from the proportion of infested plants (at least not by a linear function).

3- Whole plant infestation can be estimated from the infestation of the upper six nodes between the seventh and ninth week after planting.

References

Barral, J.M. and L.B. Zago. (1983): Programa para el manejo integrado de insectos y ácaros del algodón. Boletín Nº 71. EEA Sáenz Peña - INTA. Pp 17-29.

Cochran, W.G. (1974): Simple Random Sampling (Chap. 2) and Cluster Sampling (Chap. 9) In: Sampling Techniques. Ed. Continental S.A. Argentina.

Cochran, W.G. (1983): Planning and Analysis of Observational Studies. John Wiley and Sons. USA. Pp 50-73.

Elberson, L.R. and J.B. Johnson. (1995): Population trends and comparison of sampling methods for *Diuraphis noxia* and other cereal aphids (Homoptera: Aphididae) in northern Idaho. Environmental Entomology. 24(3):538-549.

Hardee, D.D., M.T. Smith, A.A. Weathersbee and J. Snodgrass. (1994): Sampling of the cotton aphid (Homoptera: Aphididae) in cotton. Southwestern Entomologist: 19(1):33-44.

Kapatos, E.T., E.T. Stratopoulou, A. Sahinoglou and J.A. Tsitsipis. (1996): Development of an optimum sampling plan for the population of *Aphis gossypii* (Hom, Aphididae) on cotton in Greece. J. Applied Entomology: 120(4):245-248.

Pielou, E.C. (1977): Spatial patterns in one-species populations. (Chap 11) In: Mathematical Ecology. John Wiley and Sons. Canada.

Rogers, A. (1974): Statistical Analysis of Spatial Dispersion. Pion Limited. London. Pp 1-52.

Scheaffer, R.L., W. Mendenhall and L. Ott. (1979): Elementary Survey Sampling. Duxbury Press. USA. pp 5-7; 141-161.

Steel, R.G. and J.H. Torrie. (1988): Bioestadística: Principios y Procedimientos. 2da Edición (1a en español). Mc Graw Hill. México. Pp 7-35; 83-116.

Southwood, T.R.E. (1995): Ecological Methods. with Particular Reference to the Study of Insects Population. 2nd Edition. Chapman and Hall. USA. Pp 7-47.

Williams, M.R., T.L. Wagner and J.L. Willers. (1995): Revised protocol for Scouting Arthropod Pests of Cotton in the Midsouth. Technical Bulletin 206. Mississippi Agricultural and Forestry Experimentation Station. 28 pp.

Table 1. Descriptive statistics of aphids infestation in the cotton crop cycle, lot A (direct seeding over wheat) and lot B (conventional seeding).

Date	Lot A: Seeding over wheat						Lot B: Conventional seeding					
	Aphids per plant			Proportion infested			Aphids per plant			Proportion infested		
	M	SD (S)	SD(C)	M	SD (S)	SD (C)	M	SD (S)	SD(C)	M	SD (S)	SD (C)
27-28/12/95	5.38	83.70	2.32	0.33	0.166	0.071	41.22	361.56	9.17	0.862	0.141	0.087
4-5/1/96	1.82	19.20	0.68	0.27	0.161	0.045	30.82	428.46	11.88	0.954	0.112	0.070
11-12/1/96	4.05	56.88	1.79	0.18	0.150	0.045	35.59	284.37	7.89	0.123	0.140	0.124
18-19/1/96	181.30	1354.23	69.07	0.98	0.087	0.063	317.45	2917.22	80.90	0.762	0.158	0.139
1/2/96	56.04	1117.87	34.25	0.89	0.134	0.045	--------	--------	--------	--------	--------	--------
8-9/2/96	17.15	167.18	6.47	0.93	0.123	0.032	118.78	972.98	37.93	0.893	0.134	0.061
14-15/2/96	16.00	75.83	4.70	0.96	0.105	0.014	36.13	198.01	0.51	0.952	0.112	0.006
21-22/2/96	10.92	41.68	2.84	0.98	0.095	0.025	8.91	40.46	0.26	0.924	0.125	0.019
27/2/96	--------	--------	--------	--------	--------	--------	5.09	33.69	0.19	0.655	0.167	0.057
5-7/3/96	1.78	21.64	0.73	0.40	0.169	0.027	1.54	10.49	0.10	0.431	0.170	0.032
11-14/3/96	0.11	1.51	0.05	0.07	0.123	0.017	0.67	7.97	0.07	0.190	0.152	0.033
19/3/96	3.11	31.78	1.16	0.60	0.170	0.027	--------	--------	--------	--------	--------	--------
General	31.19	122.55	11.28	0.60	0.135	0.037	46.43	144.23	14.89	0.675	0.166	0.063

M: mean.
SD(S): standard deviation, simple units sample (S).
SD(C): standard deviation, cluster sample (C).

Figure 1. Bound of error of estimation evolution with number of aphids by plant, lot A.

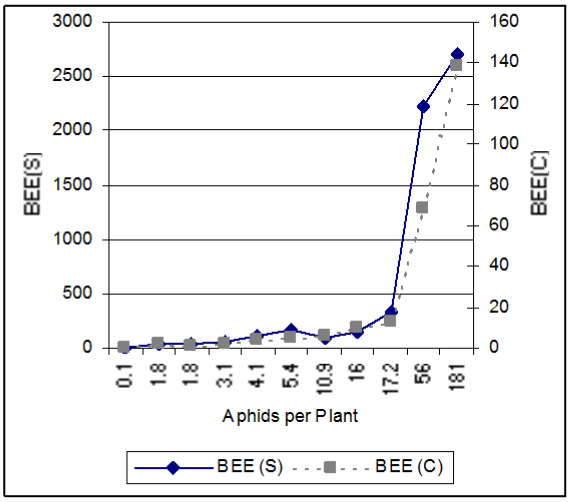

Figure 2. Bound of error of estimation evolution with number of aphids by plant, lot B.

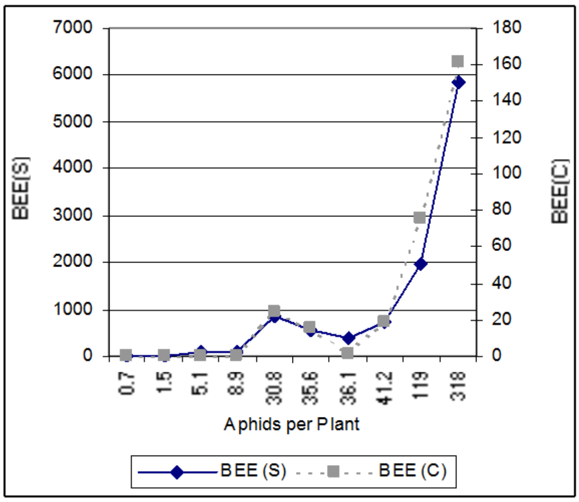

Figure 3. Bound of error of estimation evolution with proportion of infested plants, lot A.

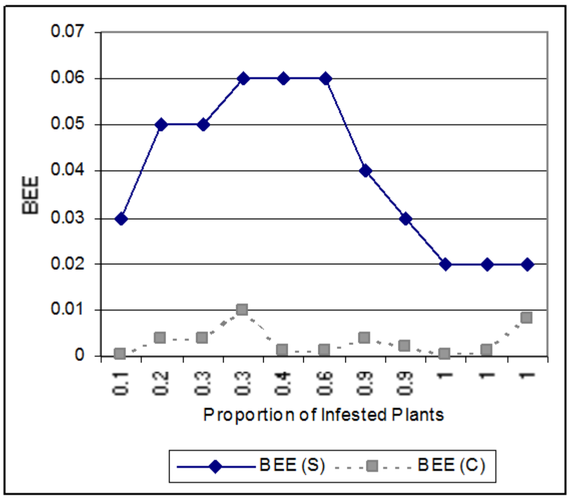

Figure 4. Bound of error of estimation evolution with proportion of infested plants, lot B.

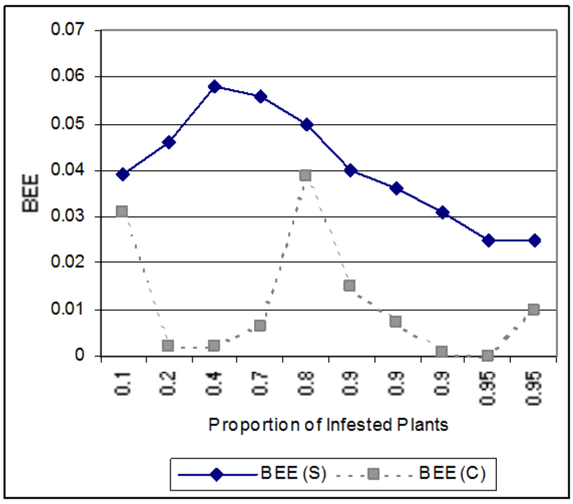

Figure 5. Relation between number of aphids by plant and proportion of infested plants, lot A.

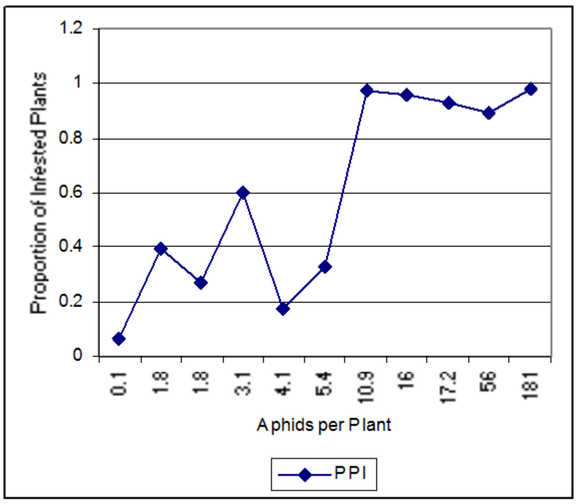

Figure 6. Relation between number of aphids by plant and proportion of infested plants, lot B.