Distribution parameters and sampling of *Brachycaudus helichrysi* in sunflower fields

Isabelle BADENHAUSSER
INRA, Laboratoire de zoologie, 86600 Lusignan (France)

Abstract

Presence-absence notation is not possible to estimate the population mean of the leaf-curling plum aphid on its secondary host, the sunflower. A sequential procedure, based on a relationship between population mean and variance according to Taylor's power law, is proposed for estimating the population mean with a fixed level of precision. The validity of the procedure is discussed using data from field and simulations of the sequential plan. It is shown that when the relative precision is better or equal 0.20, population mean estimates are unbiased and leads to the required precision. The mean sample size is compatible with agronomical practice for this level of precision.

Key-words

Brachycaudus helichrysi, Helianthus annuus, Spatial pattern, Sequential sampling

Introduction

The leaf-curling plum aphid, *Brachycaudus helichrysi* has become a major pest of sunflower in Europe (Bujaki 1984, Badenhausser et al. 1988, Camprag et al. 1988). This aphid is better known as a pest of plum, its primary host, but is also found during spring on some cultivated herbaceous plants. On sunflower, the aphid causes the leaves to shrivel and curl not because of virus transmission but through the action of its saliva (Schmelzer 1970). We have established (Lerin & Badenhausser 1995) that the effects of *B.helichrysi* infestations are deleterious when infestations begin during the very early stages of the plants (one to two leaves) and reach >100 aphids per plant at the budding stage. If the infestations decrease before or at this stage, the yield loss is negligible. According to these results systematic pest control should be avoided and should be based on population dynamics surveys, i.e., on a precise population assessment. Sampling techniques have to be established for this purpose.

The purpose of this work is i) to study if there is a relationship between the proportion of infested plants and the population mean. Such a relationship could be of prime interest because counting aphids is very tedious; ii) to model the relationship between mean and variance of counts according to Taylor's power law and iii) to propose sequential sampling plans to estimate mean density for *B.helichrysi* populations. We also present the results of validation studies for the sequential plan using data sets which are not used to establish the spatial model.

Material and Methods

a) Sample data

Data are collected in The West of France over a 6-y period from 1985 to 1992. The sunflower fields which are observed range in size from 1 to 2 ha and are grown following standard agronomic practices. Sunflower cultivars are Mirasol, Viki, Isomax, Albena, Vidoc and Frankasol. Fields are naturally infested with *B.helichrysi* populations and sampling consists in counting visually the aphids and the growth stage of the plant. Sampling is carried out during the whole period of infestation, i.e., from the beginning of the infestations to their natural decline due to biological control. Sample sizes range between 50 to 640 plants, with a sample unit consisting in 2 adjacent plants on sunflower sowing row. This sample unit is used because neighboring plants are infested independently of each other, leading to a smaller variance than single plants (Badenhausser 1994). 60 independent data sets are then obtained.

The relationship between the proportion of infested plants and population mean is established using data for which we have individual counts per plant (and not per sample unit of 2 plants), i.e., 31 data sets.

The relationship between mean and variance of the number of aphids per 2 adjacent plants is calculated using 55 data sets, while the 5 other data sets are used to validate the sequential plan. They have been chosen because their means ranged from 5.8 to 188.2 aphids per sampling unit , i.e., low to high levels of aphid infestations (Figure 2, Data sets A, B, C, D, E).

b) Relationship between mean and variance

Taylor's power law (1961) describes a relationship between mean and variance of counts per unit. This is an empirical model which fits well a great number of species (Taylor et al. 1980):

$$\log_{10} s^2 = \log_{10} a + b \log_{10} \bar{x}$$

where s^2 and $\bar{x}$ are respectively the variance and the mean of the counts per sample unit (sum of the counts on 2 adjacent plants), a is a sampling factor and b a specific measure of aggregation. Parameters are estimated using least square regression coefficients.

We calculate for each sample (55 data sets), mean and variance of the number of aphids per unit of 2 adjacent plants and the relationship is established for the 2 main cultivars Viki and Mirasol, all years pooled. The homogeneity of the slopes of the regressions for the two cultivars is tested through an analysis of variance. A global relationship is established all cultivars and years pooled and the slope of the regression is compared to that of a random distribution, i.e., to 1 (Taylor et al. 1980).

c) Sequential sampling plan

Sequential sampling can be used to estimate mean population density (Kuno 1969) or to classify populations as requiring or not pest management action (Wald 1947). The sample size is flexible and depends upon the values of the observations taken. It achieves a compromise between acceptable levels of precision or risk levels, and a sample size as small as possible. Sequential sampling can be used when random sampling is chosen, i.e., when sample units are infested independently of each other. This has been shown for *B. helichrysi* populations (Badenhausser 1994) for a sample unit consisting in 2 adjacent plants.

We develop sequential sampling to obtain population estimates with a fixed level of precision because it could be used as well as in pest management as in population dynamics studies. Green (1970) developed such sequential sampling plan using Taylor's power law to describe the variance-mean relationship. The criteria to stop or to continue sampling is given by the equation:

$$T_n^{b-2} = \frac{D_0^2}{a n^{1-b}}$$

where T_n is the cumulative number of aphids for the n sample units, D_0 is the relative required precision (in per cent of the mean, i.e., for example 0.20), a and b the parameters of Taylor's power law as defined previously. This leads to a table or a graph of the cumulative number of insects counted on the observed sample units against the number of sample units. The precisions 0.10, 0.20 and 0.30 are used to develop the sequential plans.

The validation of the sequential plan is carried out using 5 data sets which have not been used to establish the variance/mean relationship. The validation consists in simulating the sequential sampling plan. Units are randomly chosen in the data set until the stopline is reached. Then we can calculate for each such simulated run :

- the estimate of the mean $\bar{x} = Tn/n$;

- the precision really obtained : $D = s / \bar{x}\sqrt{n}$, with s the standard error of the counts for the sample, $\bar{x}$ the sample mean and n the final sample size.

We simulate the sequential plan 100 times per data set and for each simulated plan.

Analyses are carried out using Splus statistical language (Becker et al. 1988).

Results and discussion

a) Relationship between the percentage of infested plants and the population mean

The relationship between the percentage of infested plants and the population mean shows (Figure 1) that the percentage of infested plants increases very rapidly with aphid numbers. When population mean reaches less than 20 aphids per plant, 100% of the plants are infested. For lower means, binomial sampling could be based on presence-absence notation. But, as a damaging effect is observed when populations reach more than 100 aphids per plant, it is not possible to use presence-absence notation to take pest management decision.

b) Relationship between mean and variance

Figure 2 shows the relationship for the whole data, and for cultivar Viki (19 data sets) and Mirasol (20 data sets). There is no difference for b between the two main cultivars : b=1.532 for Viki and b=1.525 for Mirasol (F=2.23, P=0.14, df=1,36). That leads to the hypothesis that there is no cultivar effect on the relationship. Taylor's power law provides an adequate description of variance/mean relationship with R^2=0.96, a=3.554 and b=1.528 with $SE(b)$=0.043 for the whole data. Index b is significantly >1 indicating clearly that *B.helichrysi* populations are aggregated. Such result is not surprising because aphids reproduce themselves by parthenogenesis. The effect of the year on the relationship is not studied because insufficient replicates are available. However, the range of densities represented in our data is large (from 0.1 to 260.2 aphids per sample unit) (Figure 2), so that it facilitates accurate estimates of the power law parameters (Taylor 1984). Sequential sampling plan is then calculated using the parameter estimates of the pooled data.

c) Sequential sampling plan

The sequential sampling plan is developed using Green's formula (1970) with Taylor's parameters : a=3.554 and b=1.528 (Figure 3). The practical use of the graph can be illustrated with the following random sequence of counts as an example:

14	12	4	6	1	5	10	5	19	7	0
14	8	12	2	5	13	1	3	25	12	40
0	7	31	4	10	6	3	11	3	0	18
27	...									

If we set to 3 the minimum sample size we obtain thus for n=3 T_n=30. This point is below the stopline whatever the level of precision: the precision that we have chosen is not reached, and we need to observe one more sample unit; for n=4 we obtain T_n=36 aphids which is again below the stopline; for n=5 we have T_n=37 which is again below the stopline, etc until n=15 units with T_n=119 aphids: this point is above the stopline for the level of precision D_0=0.30. This means that the estimate of the mean by T_n/n=119/15=7.93 aphids per sample unit leads to a relative precision of 30%. To reach higher precision, for example D_0=0.20 we need to observe n=30 sample units, i.e., T_n=300 aphids. This point is above the stopline for D_0=0.20, and the mean T_n =300/30=10 aphids per unit is estimated with the precision level D_0=0.20.

The sequential plan that we have established requires that the sample size is large when the level of precision D_0=0.10 : n>50 units, i.e., 100 plants. For research purpose, this sample size could be realistic but it is generally incompatible with agronomic practice or with pest management purpose. In that case, sampling 20 units is the upper limit especially when counts are high. The precision level D_0=0.20 leads to a sampling plan which may be more realistic under current practices. Means <1 aphids per unit require a sample size> 90 units. Means from 1 to 5 aphids per unit require from 90 to 40 units and means from 5 to 20 aphids per unit require from 40 to 20 units. When means are higher, the sample size decreases slowly. For means near the action threshold, the sample size is 11 units for D_0=0.20.

d) Validation of the sequential procedure

Average results of the simulation procedures carried out on the 5 data sets show (Table 1) that for data set A the true levels of precision obtained are in average worse than the desired levels of precision, especially when the required precision is D_0=0.30. This is not the case for the other data sets. This could be explained by the fact that the point (log x, log $s2$) for this data set is not well represented on the variance-mean relationship (Figure 2).

The average means established from the 100 simulations are quite near the 'true' means (Table 1) with a tendancy to overestimate the means, whatever the data sets.

Mean sample sizes are reasonable for the relative precision D_0=0.20 which requires from 8 to 39 units for our data sets. This precision is preconised by Southwood (1978) for pest management purposes.

The results that we have presented concern the average criteria of the sequential sampling plan. It is important to underline that the variability of these criteria is important. As an example, for data set C (intermediate density: 49.8 aphids per unit) mean estimates ranged from 42.4 to 56.8 aphids per unit when D_0=0.10 for the simulated sequential sampling plans. When the level of precision is D_0=0.30, means ranged from 20.9 to 91.0 aphids per unit. This is also observed for the 2 other criteria (Badenhausser, Submitted). Such results were also mentioned by Hutchison et al. (1988). They are explained by the deviation of the data from the spatial model on which the plan is based and by the sequential process itself (Kuno 1972). They prove that the simulation step is necessary to select plans which provide a compromise between minimum sample size, accurate estimates of population density and a real control of the level of precision.

Conclusion

The sequential plan we put forward, based on power law parameters, seems to be of practical importance because it allows population mean estimates with a sample size which is not too large and with a relative level of precision (0.20) which is reasonable for pest management purposes.

Acknowledgments

The research was supported in part by Centre Interprofessionnel des Oléagineux Métropolitains.

References

Badenhausser, I. 1994. Spatial patterns of alate and apterous morphs of the leaf-curling plum aphid (Homoptera: Aphididae) in sunflower fields. Environ. Entomol. 23: 1381-1390.

Badenhausser, I. Sequential sampling of *Brachycaudus helichrysi* (Homoptera: Aphididae) in sunflower fields. Submitted to J. Econ. Entomol.

Badenhausser, I., J. Lerin & R. Bournoville. 1988. Studies on yield losses caused by two sucking pests of sunflower in France: *Brachycaudus helichrysi* and *Lygus spp.* Proceedings of the 12 th International sunflower conference, Novi Sad, Yougoslavia, 2: 158-163.

Becker, R.A., J.M. Chambers & A.R. Wilks. 1988. The newS language. A programming environment for data analysis and graphics. 702 pp., Wadsworth & Brooks/cole advanced Books & software, Pacific Grove, California.

Bujaki, G. 1984. Study of aphids damaging on sunflower plants in many regions of Hungaria in the period 1979-1984. Novenytermeless 20: 533-540.

Camprag, D., R. Sekulic, R. Thalji, T. Keresi & R. Almasi. 1988. Studies of the occurrence and harmfulness of aphids (Homoptera: Aphididae) on sunflower. Proceedings of the 12th international sunflower conference, Novi Sad, Yougoslavia, 2:170-171.

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

Hutchison, W.D., D.B. Hogg, M. Ashraf, R.C. Berbert & 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.

Kuno, E. 1969. Some notes on population estimation by sequential sampling. Res. Popul. Ecol. 14: 58-73.

Kuno, E. 1972. Some notes on population estimation by sequential sampling. Res. Popul. Ecol. 14:58-73.

Lerin J. & I. Badenhausser. 1995. Influence of the leaf curling plum aphid (*Brachycaudus helichrysi*) on stem diameter, seed yield and their relationship, in sunflower. Journal of Agricultural Science, Cambridge 125: 211-221.

Schmelzer, K. 1970. The cause of curly mosaic in sunflowers. Nachrichtenbl. Dtsch. Pflanzenschutzdienstes, Braunschweig, 24: 1-5.

Southwood, T.R.E. 1978. Ecological methods, 2nd ed. Chapman & Hall, London.

Taylor, L.R. 1961. Aggregation, variance and the mean. Nature 4: 732-735.

Taylor, L.R. 1984. Assessing and interpreting the spatial distributions of insect populations. Annu. Rev. Entomol. 29: 321-357.

Taylor, L.R., I.P. Woiwod & J.N. Perry. 1980. Variance and the large scale spatial stability of aphids, moths and birds. J. Anim.Ecol. 49: 831-854.

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

Data sets	Sample mean	Confidence interval	D_0	Average statistics (100 simulations) mean	sample size	precision
A	5.8	1.1	0.10	5.8	155	0.14
			0.20	6.2	39	0.25
			0.30	6.1	18	0.36
B	188.2	29.9	0.10	193.2	30	0.10
			0.20	184.7	8	0.21
			0.30	195.2	4	0.29
C	49.8	5.4	0.10	49.9	57	0.08
			0.20	51.0	14	0.16
			0.30	51.9	7	0.22
D	6.0	1.2	0.10	6.0	128	0.10
			0.20	6.1	39	0.18
			0.30	5.9	18	0.26
E	53.2	4.9	0.10	52.1	56	0.10
			0.20	52.7	14	0.21
			0.30	54.8	7	0.29

Table 1: Average statistics for the simulation of the sequential sampling plan, at three pre-fixed levels of precision (D_0). Sample unit= 2 adjacent plants

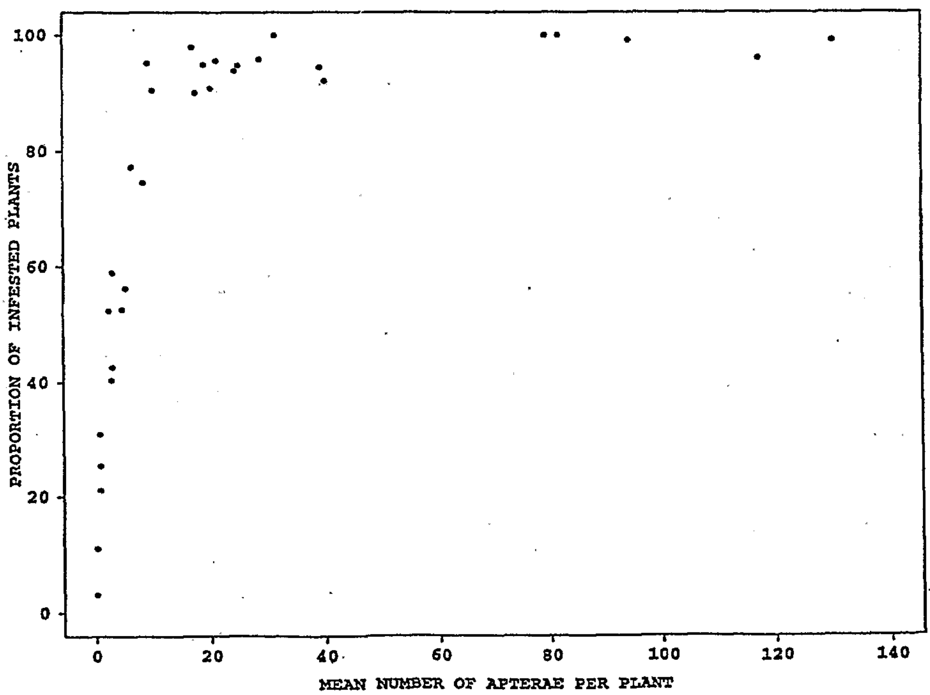

Figure 1: Relationship between the proportion of infested plants and the population mean per plant

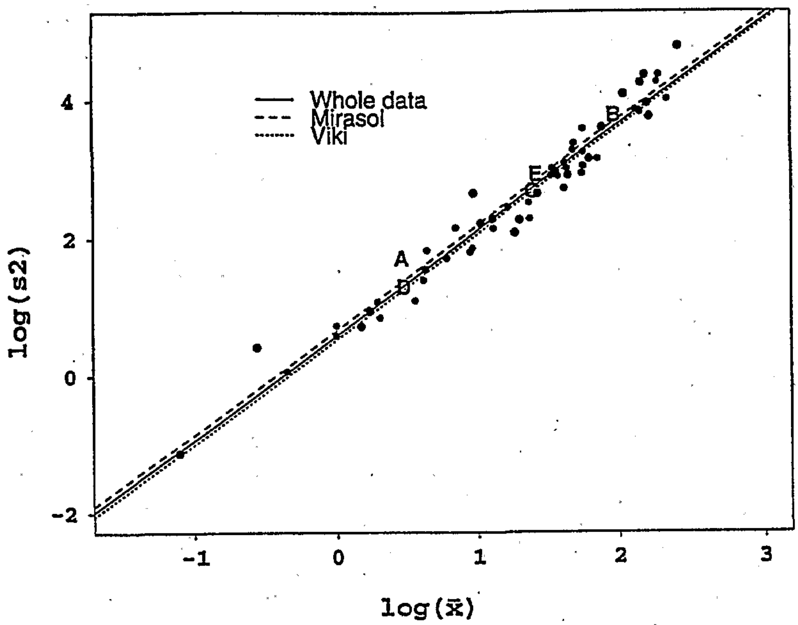

Figure 2: Relationship between mean tranformed in logarithm (log $\bar{x}$) and variance transformed in logarithm (log s2) of counts per sample unit of 2 adjacent plants, for whole data (55 data sets) and for the 2 main cultivars Viki and Mirasol. Letters identify the data sets used for the validation of the sequential sampling plan.

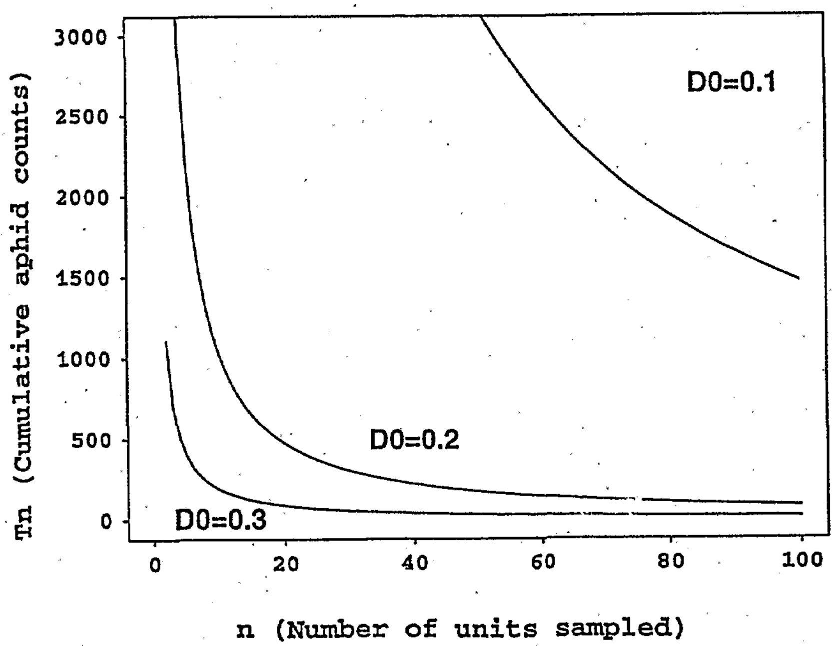

Figure 3: Sequential sampling plan for the estimate of *B. helichrysi* population mean with a fixed level of precision. Three levels of precision are represented : D_0 = 0.10, D_0 = 0.20, D_0 = 0.30.