Bulletin of Entomological Research (2002) **92**, 483–488 DOI: 10.1079/BER2002200

Spatial and temporal distribution of *Anopheles gambiae* s.l. (Diptera: Culicidae) in two Tanzanian villages: implication for designing mosquito sampling routines

E.B. Magbity[1, 2*] and J.D. Lines[2]

[1]Department of Community Health, College of Medicine and Allied Health Sciences, Fourah Bay College, PMB, Freetown, Sierra Leone: [2]Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK

Abstract

This paper describes the spatial and temporal distribution of *Anopheles gambiae* s.l. Giles in two Tanzanian villages based on data collected from a five-month intensive mosquito sampling programme and analysed using Taylor's power law. The degree of spatial aggregation of female *A. gambiae* in each village was similar to its corresponding temporal aggregation, indicating that in designing sampling routines for estimating the abundance of mosquitoes, sampling effort should be allocated equally to houses (spatial) and nights (temporal). The analysis also showed that for a given amount of sampling effort, estimates of village-level mosquito abundance are more precise when sampling is carried out in randomly selected houses, than when the same houses are used on each sampling occasion. Also, the precision of estimating parous rates does not depend on whether mosquito sampling is carried out in the same or a random selection of houses. The implications of these findings for designing sampling routines for entomological evaluation of vector control trials are discussed.

Introduction

Insect distributions are generally aggregated, that is, individuals are more clustered than would be expected if a random distribution applied (Southwood, 1978). This clustering which arises as a compounding of biological and environmental factors, sampling variation and sampling bias often gives rise to problems of precision in the evaluation of vector control interventions. Such interventions normally aim to reduce vector densities and to establish this, it is necessary to estimate and compare mean densities in areas with and without intervention. In the case of malaria vector control trials, this typically entails estimating in each of a series of villages the mean vector density for the whole village over at least a year (Magesa *et al.*, 1991; Lindsay *et al.*, 1993). However, mosquito sampling is carried out at house-level but the densities are estimated at village-level. Hence, the greater the clustering of mosquitoes, the more difficult it is to obtain adequately precise village-level estimates based on a representative sample of houses. In practice, even substantial absolute differences in observed village-level mean densities often cannot be shown to be statistically significant (Magbity, 1999). Moreover, even sampling over a one year period may not be sufficient because cyclical fluctuations in vector abundance are likely to confound the assumption that reduction in vector abundance is a result of the control measure.

In order to increase precision and to reduce sampling

*Fax: +(232) (22) 229214
E-mail: magbity@yahoo.com

error, it is necessary to increase sampling effort, but in practice available resources invariably limit sampling efforts. Often, designing an entomological sampling routine for vector control trials involves making *ad hoc* decisions about how to apply the resources available for vector sampling in a manner that will maximize the precision of the estimates obtained. This is rarely a simple task, since variation in both space and time must be considered and there are no guidelines for making these decisions.

The question therefore arises of how best to distribute sampling effort in order to maximize the precision of the estimates of village-level means. Suppose, for example, that monthly estimates of mean density in each village are required, and that no more than 12 light-trap samples can be carried out in each village, monthly. Is it better to trap six times in each of two houses, or twice in each of six houses, or four times in each of three houses? And in the latter case, is it necessary that the nights should be at weekly intervals, or would the easier task of sampling over four consecutive nights yield a similar amount of information? Should the same 'fixed' houses be sampled on each occasion, or should a new set be chosen randomly on each occasion? A further complication arises when, as is normally the case, estimates of the parous rate and sporozoite rate are also required, in addition to estimates of density. Is the best routine for comparing densities between villages also the best for comparing parous rates?

Smith *et al*. (1995) started to tackle the sampling by applying Bayesian techniques to map densities of *Anopheles gambiae* Giles and *Anopheles funestus* Giles (Diptera: Culicidae) in a Tanzanian village. They predicted mosquito densities based on a model that estimated the degree of variation attributed to various spatial and temporal factors. The present study tackles a similar problem but using a more direct approach, and hopefully will provide the basis for developing specific rationales for distributing sampling efforts when estimating village-level mosquito abundance and parous rates in vector control trials. The aim was to sample mosquitoes in an unusually intense manner in both space and time and then compare spatial and temporal variation. In addition, this intensive sampling regime allowed the use of random sub-samples of the data to simulate what the result would have been for various alternative strategies of less intense sampling.

Materials and methods

Study area

Data for this study were collected from two hamlets, Enzi Mnundu and Tengeni Central, situated in the Muheza District (5° 10′N, 38° 47′E) of Tanzania. This region had two rainy seasons a year; a short rainy season from December to January, and a long rainy season from April to June.

Falciparum malaria was holoendemic with *A. gambiae* and *A. funestus* the primary malaria vectors. These vectors bred mainly in the swamps and numerous ditches within and around the villages.

Study design

In each hamlet, six designated houses were selected randomly and a room in each was chosen for mosquito sampling. All sleeping places in the sampling rooms were supplied with untreated bed nets.

Data were collected for 18 weeks, from February to June 1996. In each week a hamlet was randomly selected for mosquito sampling for the first set of three consecutive nights, followed by sampling in the other hamlet for the second set of three nights. On each sampling occasion, mosquitoes were sampled simultaneously in all six sampling rooms of the hamlet.

Mosquitoes were sampled using Centers for Disease Control (CDC) miniature light traps, placed beside an occupied untreated bed net as described by Lines *et al*. (1991). Prior to each sampling occasion, numbered light-traps and batteries (6V, 10Ah) were distributed randomly among the various houses. Each householder was instructed in the proper operation of the trap and participated in the study by turning the traps on at sunset. The traps were turned off in the morning by the project staff, who also recorded the trap and battery numbers used in each room, and the number of people who had slept in the room the previous night. The staff also enquired if the people noticed any malfunctioning of the trap during the night. Six data points were discarded for traps that did not work properly.

Statistical analysis

Data were analysed using: (i) Taylor's power (Taylor, 1961) to compare the magnitude of spatial and temporal aggregation of *A. gambiae* populations; and (ii) a novel bootstrapping technique to compare the relative sampling error in estimating mosquito abundance using different mosquito sampling routines.

Spatial and temporal aggregation of mosquito populations

The spatial and temporal variation of *A. gambiae* in each of the study villages were determined using Taylor's power law (Taylor, 1961), which states that the spatial variance, s^2, characteristic of a species at a particular stage in its development, is proportional to a fractional power of the mean population density, m, at that place. That is,

$$s^2 = am^b$$

The equation is typically linearized with logarithmic transformation, to

$$log_{10}(s^2) = log_{10}(a) + b\, log_{10}(m)$$

Taylor claimed that the intercept, a, is a scaling factor related to sample size, and that the slope, b, is an index of aggregation that is dependent upon species behaviour and the environment. A value of $b = 1$ indicates random distribution, while $b > 1$ indicates aggregated distribution, and $b < 1$ indicates a regular distribution. An appropriate transformation can be found from the formula $p = 1 - b/2$. According to this relationship: if $p = 0$, a logarithmic transformation is appropriate for a given set of data; if $p = 0.5$, a square root transformation is appropriate. Taylor *et al*. (1978) showed that most insect populations have b values between 1 and 2, giving transformation factors, p, between 0 and 0.5, indicating a transformation somewhere between the square root and the logarithmic.

Taylor *et al*. (1980) also postulated that just as each species has its own fixed, functional relationship between spatial variance (s_s^2) and mean population density (m_s) over an area

at all times described by a power law, temporal variability (s_t^2) is also a power function of mean population density (m_t) over time at all places, given as

$$s_t^2 = am_t^b$$

or

$$log\ s_t^2 = log(a) + b\ log(m_t).$$

For the analysis of spatial variation in mosquito abundance, the mean, x_s, and variance, s_s^2, of *A. gambiae* were calculated for each night in each village. The spatial aggregation index of mosquitoes was estimated by regression of s_s^2 against x_s, after transforming both to $\log_{10}$ scale (Taylor, 1961).

For the analysis of temporal variation in mosquito distribution, the means and variances were calculated per house per night for mosquitoes collected within each month (four weeks period). Monthly intervals were used for calculating s_t^2 and x_t in order to assess day-to-day variability within months. The variability of mosquitoes between days in a month was determined by regression of s_t^2, on x_t, after transforming both to $\log_{10}$ scale (Taylor, 1961).

All power law regression analyses of these data were carried out after excluding means, $m < 2$, and variances, $s^2 < 4$ (Taylor & Woiwod, 1982). In the present case, only one record was excluded from the data for temporal analysis, and none from the data for spatial analysis.

The least squares linear regression procedure of STATA 5.0 (Statacorp, 1995) was used to determine variables in the power law. Confidence intervals were used to determine if the slopes of the regression lines (b values) were significantly different from one.

Determining relative sampling error

Bootstrapping techniques were used to compare the relative sampling error in estimating mosquito abundance from different sub-samples of the data representing various alternative sampling routines. Only data from the first 12 weeks of sampling were used for these bootstrapping exercises, because longer periods would include larger variations in mosquito abundance. Counts of adult female *A. gambiae* caught in individual light-traps were log-transformed to the scale $\log_{10}(x+1)$. The STATA 5.0 statistical package was used to design statistical programmes capable of generating different subsets of these data, with each subset simulating possible data from a less intensive sampling routine.

For example, a programme was designed to generate a subset of the data collected from the village of Enzi by stimulating a routine of sampling one night a week in a single designated house for 12 weeks. This programme was run 1000 times to generate 1000 subsets of data. The mean log mosquito count for each of the 1000 simulated data sets was calculated, and a new data set containing the 1000 means was constructed. The mean of these means and the size of its 95% confidence interval were then calculated. The 95% confidence interval was calculated by ranking the 1000 means and excluding the top and bottom 25; the range of the remaining 950 estimates was designated the 95% confidence interval.

The relative sampling error was estimated as the percentage relative sampling error, which is expressed as: percentage relative sampling error = 50 × (size of the confidence interval) / mean (Sutherland, 1996).

This gives the error in calculating the mosquito abundance for a given sampling routine relative to the best estimate of the mean mosquito abundance.

For each sampling routine, the relative sampling error was calculated for data generated when the same houses were used, and also when the houses were randomly selected on each sampling occasion.

The procedure used for calculating the relative sampling error in estimating parous rate was similar to that described above for mosquito abundance. Parous rates were calculated for each of the 1000 simulated data sets, as a fraction of parous females to the total number of females dissected in each data set. These were used to construct a new data set containing 1000 parous rates.

Results

Power law regression analysis

Power law regression analysis for both spatial and temporal distributions of *A. gambiae* yielded slopes significantly greater than one, signifying spatial and temporal aggregation of *A. gambiae* mosquitoes (table 1 and fig. 1).

Table 1 also showed that the transformation indices (estimated from $p = 1 - b/2$) were not significantly greater than zero, justifying logarithmic transformation of data on mosquito abundance.

Precision of estimating mosquito abundance

Figure 2 shows the relative sampling error in estimating mosquito abundance as a function of sampling routine involving various amounts of sampling effort. For various

Table 1. Regression coefficients of Taylor's power law for spatial and temporal variability of *Anopheles gambiae* s.l.

	b (95% C.I.)		p (95% C.I.)		r^2	
	b_s	b_t	p_s	p_t	r_{s2}	r_{t2}
Enzi	1.99 (1.58; 2.41)	1.83 (1.55; 2.11)	0.00 (0.21; −0.20)	0.00 (0.22; −0.06)	0.63	0.84
Tengeni	1.70 (1.44; 2.01)	1.77 (1.47; 2.06)	0.15 (0.28; −0.01)	0.11 (0.26; −0.03)	0.76	0.84
Combined villages	1.79 (1.57; 2.04)	1.82 (1.63; 2.08)	0.10 (0.21; −0.02)	0.09 (0.19; −0.04)	0.72	0.86

Subscripts s and t stand for spatial and temporal values respectively; p is a transformation factor; r^2 = goodness of fit of the regression model, $p = 1 - b/2$.

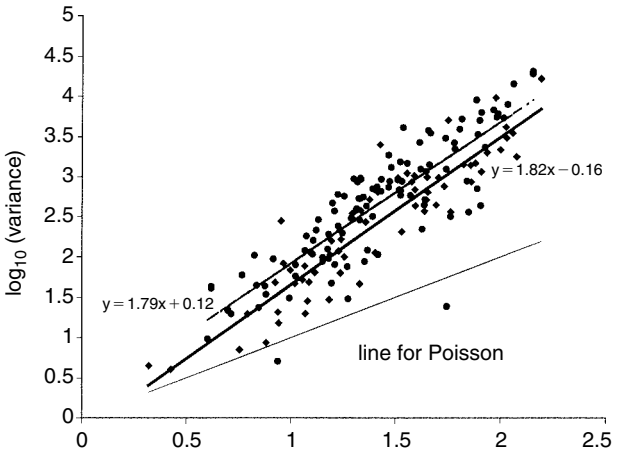

Fig. 1. Taylor's power law regression for spatial analysis of variance against the mean density of *Anopheles gambiae* in each village. Each data point represents the log-transformed mean and variance of the number of *A. gambiae* per light trap caught over all the houses sampled on a particular night in each village (● ----, spatial; ♦ ——, temporal).

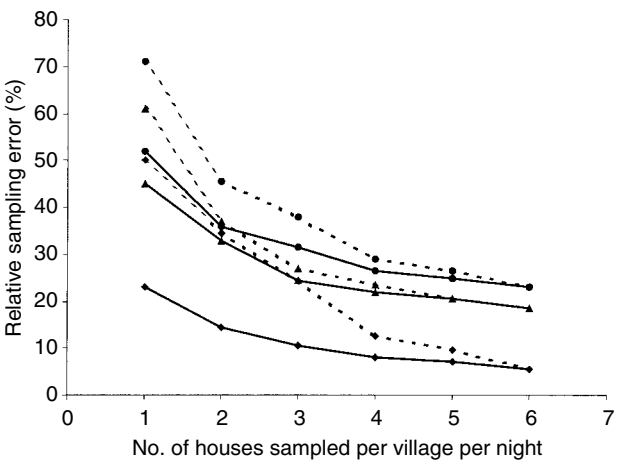

Fig. 2. Relative sampling error in estimating mosquito abundance using various sampling designs in either the same houses (---------) or in a random selection of houses (——) on each occasion (♦, weekly; ▲, fortnightly; ●, monthly).

predetermined levels of sampling efforts, the figure compares the relative sampling error in estimating mosquito abundance by sampling in the same houses to random selection of houses on each occasion. The results show that the relative sampling error in estimating mosquito abundance decreases as the number of houses sampled or the frequency of sampling increases. The results also show that for a given amount of sampling effort, the relative sampling error was lower when sampling was carried out in randomly selected houses, than when the same houses were used on each sampling occasion.

Figure 3 compares the relative sampling error in estimating mosquito abundance by sampling once a week with that obtained from sampling on two consecutive nights a fortnight, for the same total sampling effort. The figure clearly reveals that sampling at weekly intervals resulted in less error than sampling on two consecutive nights per fortnight.

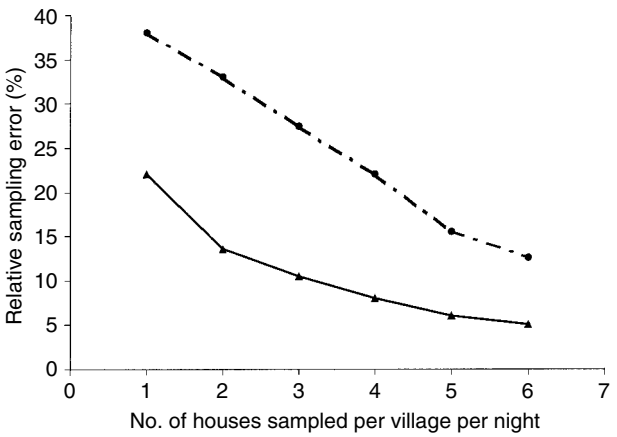

Fig. 3. Comparison between the relative sampling error in estimating mosquito abundance by sampling once a week (-▲-) and two consecutive nights a fortnight (--- ● ---), using the same total sampling effort.

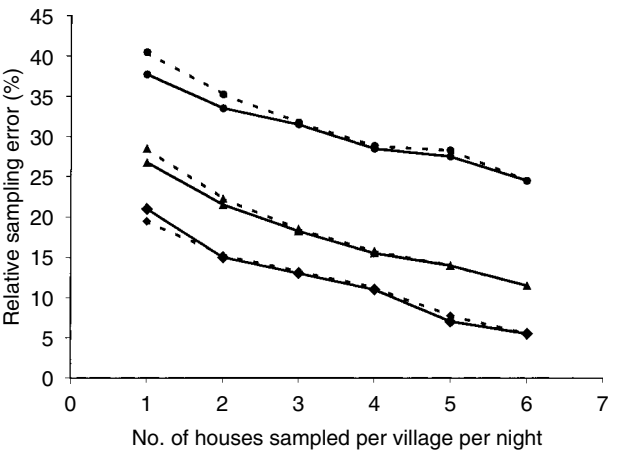

Fig. 4. Relative sampling error in estimating parous rates using various sampling designs in either the same houses (------) or in a random selection of houses (——) on each occasion (♦, weekly; ▲, fortnightly; ●, monthly).

Precision of estimating parous rates

Figure 4 compares the relative sampling error in estimating parous rates for different predetermined levels of sampling effort and sampling routines. The figure shows that the relative sampling error in estimating parous rates reduces as either the number of houses sampled or the frequency of sampling increases. However, unlike the estimation of mosquito abundance, there seemed to be no clear differences in sampling error in estimating parous rates between sampling in a random selection of houses and sampling in the same houses.

Discussion

Power law analysis of the spatial distribution of *A. gambiae* mosquitoes in the two Tanzanian villages revealed clustering. Spatial clustering indicates that more mosquitoes were found in some houses than others. This result agrees with that obtained by Ribeiro *et al*. (1996) in Ethiopia, where

clustering was observed mainly at the peripheral houses of the study village. According to their results, factors related to the location of the houses, such as their distance from breeding sites and surrounding rather than house-specific factors such as house design and presence of open eaves were responsible for the observed clustering. In the present study, the reasons for the clustering are not clear. However, since the villages were small with lots of scattered breeding sites, it is possible that both the general location of the houses and factors specific to individual houses could have been responsible for the differences observed in the number of mosquitoes caught in different houses on the same night.

When mosquito samples are taken over a period of time, the issue of temporal variability becomes crucial. Analysis of temporal variability for *A. gambiae* mosquitoes revealed a clumped within-month distribution, probably due to the effect of meteorological factors. Some of these factors (e.g. rainfall during the previous week) have long-term impact on mosquitoes and may even affect their actual abundance, while others (e.g. wind velocity) only act over the short-term and affect only the catchability of mosquitoes (Bidlingmayer, 1985). The relative degree to which these long and short-term factors contribute to the number of mosquitoes caught by any given sampling method is unknown. Hence the degree to which the number of mosquitoes caught in light traps actually represents the mosquito population is also unknown. However, by considering the suggestions made here, one may be able to increase the precision of the estimates.

It should be noted that the power law has come under considerable criticism, that the index, *b*, does not differ between species and is even inconsistent within species (Downing, 1986). Therefore, the results obtained here should be considered as only initial estimates. Moreover, whereas temporal aggregation was estimated by sampling 12 nights every 28 days, only six of the more than 50 houses in each village were used to estimate spatial aggregation. The estimate of temporal variability is therefore likely to be more reliable than the corresponding estimate of spatial variability. It is possible that had more houses been used, the magnitude of the spatial variation might have been different.

The results clearly revealed less sampling error in estimating mosquito densities when sampling was carried out in a random selection of houses than when the same houses were used on each occasion. This could be because this type of routine is more liable to account for spatial variation and therefore provide more representative estimates. On the other hand, the relative sampling error in estimating mosquito parous rates were not affected by whether a random selection of houses or a fixed set of houses were used on each occasion. This is because, even though mosquito distribution is aggregated, parous and nulliparous mosquitoes are expected to be randomly distributed within a mosquito population. There are no documented environmental factors that preferentially attract mosquitoes on the basis of their parity.

Estimates of parous rates and mosquito abundances are both subject to night to night variation but the former is less subject to house to house variation. One should therefore expect estimates of village-level parous rates to be more precise and reliable than village-level mosquito abundances. By coupling this with the fact that a reduction in parous rate (mosquito survival rate) has more impact on malaria transmission than reduction on mosquito abundance (Garrett-Jones, 1964), it might be more meaningful in situations where there are very limited resources for entomological evaluation to focus entirely on estimating parous rates and ignore mosquito abundance. In such situations, a sampling routine that permits a more reliable estimate of parous rates while ignoring mosquito abundance should involve frequent sampling in fewer houses on each occasion. The same houses, rather than a random selection of houses each time, can be used for mosquito sampling, in order to simplify the sampling regime. And, in order to maximize output, it would be appropriate to sample in houses with high mosquito abundance and to dissect all the mosquitoes caught for parous rate determination.

The conclusions so far reached are rather tentative, because mosquitoes were sampled over a small area, and for a short period. Further study involving more houses sampled over a longer period, in the same and in other areas, would be required to verify these results. There are also some obvious further steps that need to be investigated such as dominant spatial and temporal factors and effect of seasonal variation on spatial and temporal variation.

Acknowledgements

The authors are most grateful to Professor Chris Curtis and Dr Caroline Maxwell for their support during the fieldwork in Tanzania. Many thanks go to the entire staff of the National Institute of Medical Research at Muheza for their hospitality and assistance in various ways. We are grateful to the people of Tengeni and Enzi village for their permission to carry out the study in their village. Clearance for this research was given by the National Institute of Medical Research and the Tanzanian Commission for Science and Technology and was funded by the Department for International Development of UK.

References

Bidlingmayer, W.L. (1985) The measurement of adult mosquito population changes – some considerations. *Journal of the American Mosquito Control Association* **1**, 328–349.

Downing, J.A. (1986) Spatial and heterogeneity: evolved behaviour of mathematical artifact. *Nature* **323**, 255–257.

Garrett-Jones, C. (1964) Prognosis for the interuption of malaria transmission through assessment of the mosquito's vectorial capacity. *Nature* **24**, 1173–1175.

Lindsay, S.W., Alonso, P.L., Armstrong-Schellenberg, J.R.M., Hemingway, J., Adiamah, J.H., Shenton, F.C., Jawara, M. & Greenwood, B.M. (1993) A malaria control trial using insecticide-treated bed nets and targeted chemoprophylaxis in a rural area of The Gambia, West Africa. 7. Impact of permethrin-impregnated bed nets on malaria vectors. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **87** (Supplement 2), 45–51.

Lines, J.D., Curtis, C.F., Wilkes, T.J. & Njunwa, K.J. (1991) Monitoring human-biting mosquitoes (Diptera: Culicidae) in Tanzania with light-trap hung beside mosquito nets. *Bulletin of Entomological Research* **81,** 77–84.

Magbity, E.B. (1999) *Methods for entomological evaluation of insecticide treated bed net trials*. PhD thesis. University of London.

Magesa, S.M., Wilkes, T.J., Mnzava, K.J., Njunwa, K.J., Myamba, J., Kivuyo, M.D.P., Hill, N., Lines, J.D. & Curtis, C.F. (1991) Trial of pyrethroid impregnated bed nets in an

area of Tanzania holoendemic for malaria. Part 2. Effect of the malaria vector population. *Acta Tropica* **49**, 97–108.

Ribeiro, J.M.C., Seulu, F., Aboe, T., Kidane, G. & Teklehaimanot, A. (1996) Temporal and spatial distribution of anopheline mosquitoes in an Ethopian village: implications for malaria control. *Bulletin of the World Health Organization* **74**, 299–305.

Smith, T., Charlwood, J.D., Takken, W., Tanner, M. & Spiegelhater, D.J. (1995) Mapping the density of malaria vectors within a single village. *Acta Tropica* **59**, 1–18.

Southwood, T.R.E. (1978) *Ecological methods with particular reference to the study of insect populations*. 524 pp. London, Chapman & Hall.

StataCorp (1995) *Stata statistical software*: Release 5.0, College Station, Texas, Stata Corporation.

Sutherland, W.J. (1996) *Ecological census techniques*. 336 pp. Cambridge, Cambridge University Press.

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

Taylor, L.R. & Woiwod, I.P. (1982) Comparative synoptic dynamics. 1. Relationship between inter- and intra specific spatial and temporal variance/mean population parameter. *Journal of Animal Ecology* **51**, 879–906

Taylor, L.R., Woiwod, I.P. & Perry, J. (1978) The density dependence of spatial behaviour and rarity of randomness. *Journal of Animal Ecology* **47**, 385–406.

Taylor, L.R., Woiwod, I.P. & Perry, J.N. (1980) Variance and the large scale spatial stability of aphids, moths and birds. *Journal of Animal Ecology* **49**, 831–854.

(Accepted 15 August 2002)
© CAB *International*, 2002