*Ann. Zool. Fennici 34: 47–57*                ISSN 0003-455X
Helsinki 4 June 1997

# Sampling error in measuring temporal density variability in animal populations and communities

Mikko Mönkkönen & Jouni Aspi

*Mönkkönen, M. & Aspi, J., Department of Biology, University of Oulu, P.O. Box 333, FIN-90571 Oulu, Finland*

*Received 14 October 1997, accepted 18 November 1997*

Variation in animal numbers is an interesting ecological problem in both theoretical and applied research. Recent research has shown that there is a myriad of problems involved in measuring variability in animal populations, which have not been addressed in most empirical studies on fluctuations of animal densities. Therefore, we actually know less about variability in animal populations and communities than we think. It is rarely possible to accurately sample entire populations and our estimates of variability usually come from spatially restricted samples of counts drawn from local populations. The observed variability not only reflects variability in population density or size but also involves a sampling variance component. Sampling variance occurs principally due to inexactness of the counts (i.e. all individuals present in the sampling unit do not enter into samples) and spatial variance (the size of the sampling unit is inadequate to capture the dispersion pattern of individuals in the field). Many samples are affected by both of these sources of error and in most cases we are unable to separate their effects. Sampling variance usually affects the variability estimates and particularly besets small samples. When comparisons are to be made in temporal variability between communities, species, populations or sites, great care must be taken to mitigate the effects of sampling variance. If the counts are replicated in space or time then sampling error can fairly simply be estimated and removed. Even in the absence of replication, statistical methods exist allowing estimation of the sampling variance. These methods are only applicable if we are prepared to make assumptions about the distributions of the counts. We exemplify one of these methods by considering a classical case of latitudinal gradients in density variability in animal communities. We finally discuss recent studies, the results of which might be artefacts arising from sampling variance.

## 1. Introduction

Animal abundances fluctuate in time. Some taxa seem to vary more than others (Schoener 1986, Hanski 1990). A number of within-species patterns in the temporal variability of abundances have also been postulated. The magnitude of variability has been found to correlate negatively with mean population density (Järvinen 1979, Taylor & Woiwod 1980, Curnutt *et al.* 1996), body size (Gaston 1988), geographic range size (Gaston 1988) and feeding specificity (Redfearn & Pimm 1988), and positively with population growth rate (Pimm 1991, Hanski & Woiwod 1993), latitude (Järvinen 1979) and extinction probability (Pimm 1993). These temporal variations are of substantial interest to animal ecologists also because these dynamics have important, both theoretical and applied, bearings on population and community ecology as well as on conservation biology. At the community level, variation may lead to interruptions or reversals of biological interactions thereby slowing down the process of competitive exclusion or even preventing it (e.g. Chesson 1986).

However, recent research has shown that there are serious problems involved in comparisons of population variability between different species and populations (McArdle *et al.* 1990, McArdle & Gaston 1992, 1993, 1995, Gaston & McArdle 1993, 1994, Leps 1993, Link *et al.* 1994, Link & Nichols 1994, Stewart-Oaten *et al.* 1995) suggesting that we actually know less about variability in animal populations and communities than we think. McArdle *et al.* (1990) asserted that "problems associated with the measurement and interpretation of population density variability … have confounded most, if not all, previous studies of the subject". These problems are both conceptual and statistical.

Most studies of variability are based on counts of animal populations. Exact censuses of total populations are usually impossible. Therefore, we often do not know the true number of animals in a population, and hence the true density or the true variability of the density. We can only estimate these values based on samples taken from a limited area (field, forest tract, quadrate, circular sampling area etc.). Sampling errors, that is, the inexactness of the counts (Link *et al.* 1994) and spatial variance (Horne & Schneider 1995, Stewart-Oaten *et al.* 1995), usually affect variability in samples (*see also* McArdle & Gaston 1993). These sources of variability together are often called sampling variance. This component of variability is not biologically relevant, and should be removed from measures of variability. In this paper, we provide a survey of the sources of sampling variance (sampling error) and review the methods proposed to remove their effects. It is obvious that variability in multi-species assemblages involves the same problems as at the population level, and the effects of sampling variance may be quite dramatic. We, therefore, provide a look at the variability in the total density of community samples, and also briefly refer to variability in other community properties such as species number and diversity. Finally, we provide a few examples to illustrate the effects of sampling variance in real data.

## 2. Measures of variability

A large number of indices have been proposed as measures of temporal variability of animal abundances (Gaston & McArdle 1994). Of the three most commonly used, standard deviation of log-transformed densities (SDlog($X_t$), where $X_t$ is the number of individuals in the sample at time $t$), and coefficient of variation of untransformed densities (CV($X_t$) = SD$X_t$/$x$, where $x$ is the mean of $X_t$'s) provide a constant measure of variability, whereas the third, SDlog($X_t$+1) is biased (Anscombe 1948, McArdle *et al.* 1990, McArdle & Gaston 1995).

It was shown that SDlog($X_t$) and CV($X_t$) are very similar, so that if two populations have equal CV($X_t$) then the values of SDlog($X_t$) tend also be very similar, and if CV($X_t$) differs so will SDlog($X_t$) (Gaston & McArdle 1994). At rather low densities, some or even most $X_t$ values will be zero and log($X_t$) remains undefined but CV($X_t$) can still be used.

A consistent observation in ecology is that variability in population abundance of a species and average population densities are related in both space and time. There are several models to describe variance-mean relationship (*see* Gaston & McArdle 1994). A conventional way is to use Taylor's power function (Taylor 1961):

$$\log(V) = a + b\log(x), \qquad (1)$$

where $V$ is sample variance, $x$ is average abundance, and $a$ and $b$ are constants. Mean-variance plots on a logarithmic scale are often called Taylor power plots. Note that variability is independent of mean abundance when the slope of the Taylor power plot $b = 2$ (Hanski 1982, McArdle *et al.* 1990). In such a case, there is no slope between $\mathrm{CV}(X_t)$ and average abundance but the slope between $\mathrm{SDlog}(X_t)$ and average abundance is 1.

## 3. Sampling variance

### 3.1. Density variability and mean abundance

In measuring population variability, one must distinguish between variability of a population and that of a site (McArdle & Gaston 1993). Usually ecologists measure variation in the number of individuals in a spatially restricted sample from an open population, i.e. variation in density. This measure may or may not reflect population variation. It is perfectly possible to find stable local densities, for example, in source habitat patches where reproduction exceeds mortality, even though total population size varies wildly because of high variability in sink habitat patches, where populations are dependent on the immigration from the source (*see* Pulliam 1988, Pulliam & Danielson 1991, for source-sink population model). Global population size may remain stable even though densities at study plots vary, e.g. in nomadic species. Unlike population variability, community variability is a site-related problem. In this paper, we explicitly deal with densities, i.e. counts in samples representing study areas clearly smaller than the species range.

The dependence of variability on mean abundance (e.g. $b \neq 2$ in Taylor power plots) makes comparisons between samples of different sizes deceptive and affects any statement based on a single estimate of variability (*see* McArdle *et al.* 1990). The decreasing $\mathrm{CV}(X_t)$ with increasing mean abundance (*see* e.g. Helle & Mönkkönen 1986) may be solely due to a sampling error because the sampling error affects small samples more than large ones (Järvinen & Lokki 1978, Taylor & Woiwod 1982, Taylor 1984, 1986, McArdle & Gaston 1993).

### 3.2. Inexactness of the counts

Quite often all individuals present on the study plot do not enter into samples taken at any given time $t$. This particularly applies to most census schemes designed to enumerate and monitor wildlife populations. If sampling efficiency is less than 100%, even completely stable populations may appear as variable; yearly samples get 'noisy' (e.g. Järvinen & Lokki 1978, Link *et al.* 1994). Any analysis treating counts as if they represented the actual population size have been assumed to lead to biased estimates of temporal variability, even if estimates of average density were accurate (e.g. Link *et al.* 1994). This is not necessarily the case. There is a possibility of decreased variability as well. We simulated the effect of adding error variability to true variation in a time series where mean abundance was 100 individuals. Both 'true' and error variability follow the Poisson distribution. The process with added error amounting to 10% of true variability decreased variance compared with true variability in 32% of cases. So, to be exact, the observed value of variance grows in general because of sampling variance, but this does not say anything about a given single case.

Sampling from a population with $N_t$ individuals can be described by random binomial process, where individuals enter into samples with probability $p_i$, the efficiency of the method used for species $i$. Expected count, $X_t$, then becomes $p_i N_t$ with variance $(1 - p_i)p_i N_t$ (*see* Järvinen & Lokki 1978, Link *et al.* 1994) which is $(1 - p_i)X_t$. This gives $\mathrm{CV}(X_t)$ as $\sqrt{(1 - p_i)/X_t}$. The error variability because of inexactness of the counts measured by $\mathrm{CV}(X_t)$ decreases with increasing $X_t$ and $p_i$. On log-log scale, the slopes of the lines between $X_t$ and $\mathrm{CV}(X_t)$ are $-0.5$ if $p_i$ remains constant for all $X_t$. This slope corresponds to $b = 1$ in Taylor power plots. In general, rare species produce Taylor power plots with $b = 1$ because their variation follows the Poisson distribution (Taylor 1984). Therefore, if sampling efficiency is less than 1, there is a chance that one could interpret such slopes as representing populations in which temporal/spatial variation is random. This may or may not be true. Even with fairly high values of $p_i$, sample sizes should be several tens of individuals to avoid high error variability, e.g. if $p_i = 0.8$, average $X_t$ should be larger than 20 to get error $\mathrm{CV}(X_t)$ smaller than 10%.

Furthermore, sampling efficiency, $p_i$, certainly varies among species and need not be a species specific constant either. It may vary with population density, for example. At high population density social stimulus from other individuals, both conspecific and heterospecific, increases territorial activity such as singing in birds (Møller 1992, Verner 1992). This may result in higher detectability at high, when compared with low, densities (Helle & Pulliainen 1983). In a similar manner, a higher density may result in a higher mobility of individuals, which, in turn, enhances sampling efficiency (probability of being counted, trapped etc.). The probability of being counted may also become reduced at high densities, e.g. because of trap saturation (Xia & Boonstra 1992). If $p_i$ increases with $N_t$, the decrease in $\mathrm{CV}(X_t)$ with mean $X_t$ may be very steep. Summarizing, if the densities are estimated rather than counted completely, perfectly stable populations are expected to yield slopes $\leq 1$ in Taylor power plots, or slopes $\leq -0.5$ in $\log(X_t)$ vs. $\log\mathrm{CV}(X_t)$ plots.

### 3.3. Spatial variance

The second form of a sampling variance involves the problem of scale. Sampling (with $p_i = 1$) of a population, which is completely uniformly dispersed and stable in time, results in no variability, but if individuals are not uniformly dispersed, a sampling variance may appear depending on the scale of a sampling plot relative to the pattern of dispersion (Wiens 1981, Stewart-Oaten 1995). As the number of individuals in populations and communities are very seldom counted completely, but rather estimated based on a spatially restricted sample from open population(s), these temporal variability estimates often contain a (usually unknown) proportion of spatial variance. Individuals in a population may appear aggregated at a certain spatial scale, but randomly or uniformly dispersed at smaller and larger scales (Pielou 1977). Increasing the spatial area of a sampling unit inevitably increases sample size but may also affect variability independently of mean abundance. The spatial scale affects the estimate of temporal variability (Wiens 1981, McArdle & Gaston 1992) potentially confounding comparisons between different populations and species if sampling units are of a different size relative to individual territories/home ranges.

To exemplify how this form of a sampling variance affects the estimate of $\mathrm{CV}(X_t)$ in the density of a species, we calculated values of $\mathrm{CV}(X_t)$ for samples drawn from the negative binomial (representing aggregated dispersion of individuals on a larger scale) and Poisson (random dispersion) distributions, as well as from a more uniformly dispersed population assuming the spatial variance to be 0.5 times the mean abundance. In all cases, $\mathrm{CV}(X_t)$, due to spatial variance alone, clearly decreased with sample size (Fig. 1). Quite expectedly, samples forming aggregated distributions yield higher error variability than samples from random or uniform distributions which both produced slopes of $-0.5$ (i.e. $b = 1$ in Taylor power plots). If it is generally true, as suggested by Taylor *et al.* (1978), Rosewell *et al.* (1990) and Hanski *et al.* (1993) among others, that in most populations the dispersion of individuals in the field is clumped, spatially inadequate samples may produce significantly biased estimates of population variability, even with very high sample sizes, if the sampling unit was too small to control for this spatial pattern.

The situation may, however, be more complicated than shown in Fig. 1 because the dispersion pattern may vary with density. At the scale of a uniform sampling area, dispersion is more likely to be uniform at a high density than at a low one. In very dense local populations, most or all suitable sites may be occupied. Territorial behavior (non-overlapping home ranges) leads to uniform dispersion in a lower relative density as compared with non-territorial species. In less dense populations, there is space for random and aggregated dispersion. If one compares populations or species with largely unequal densities, there is a risk for even more strongly biased conclusions than shown in Fig. 1 as dispersion pattern may change from aggregated to random and/or to uniform as density increases. If clumpedness decreases (i.e. the $k$ parameter of the negative binomial increases) with abundance, slopes less than 2 in Taylor power plots result in other things being equal. Note, however, that at some other spatial scale larger than that of the uniform sampling area, increased density may lead to increased aggregation since all animals may congregate in the only available parts of the landscape that are suitable.

### 3.4. Sampling variance in community variability

Summed abundances of several species, i.e. community samples, are similarly affected by a sampling variance. As already shown for populations, due to inaccurate sampling ($p_i < 1$), error variability can be approximated by the Poisson distribution. The sums of counts from independent Poisson distributions follow this distribution pattern themselves (McArdle & Gaston 1995). Therefore, assuming that species vary independently from each other, we can use the sum of variances of individual species to estimate variance in a community sample. In other words, if we had a method to estimate sampling variance in population counts then we would also have an estimate of sampling variance in a community sample in the absence of interactions.

Community level sample sizes are often very large, hundreds of individuals, as compared with single population samples, and the bias caused by sampling variance is, therefore, obviously smaller. For example, if $p_i = 0.8$ for all species and $\sum X_{t\cdot} = 100$, CV($\sum X_t$) then becomes 4.5%. It would, nevertheless, be unwise to discount the influence of sampling variance in community samples because, even in large samples, error variability decreases with total abundance; if $\sum X_{t\cdot} = 500$ and $p_i = 0.8$, CV($\sum X_t$) is 2%. Comparisons between samples of different sizes may still, therefore, be quite misleading.

Järvinen and Lokki (1978) assessed the effect of the inexactness of counts on the variability in species richness and diversity (diversity indices) in bird community samples. They showed that in a typical single visit census of birds some 80 to 100% of the species are observed, and that the variability in species number decreases with increasing census efficiency, $p_i$. Likewise, the variability in species diversity varies with census efficiency, and particularly in species-poor communities, variability is affected if $p_i$ is relatively low. Because the sum of species variance equals the variance of the sum of counts, there is, counterintuitively, no *a priori* reason to expect that communities with many rare species would show higher variability than communities with a more even abundance distribution.

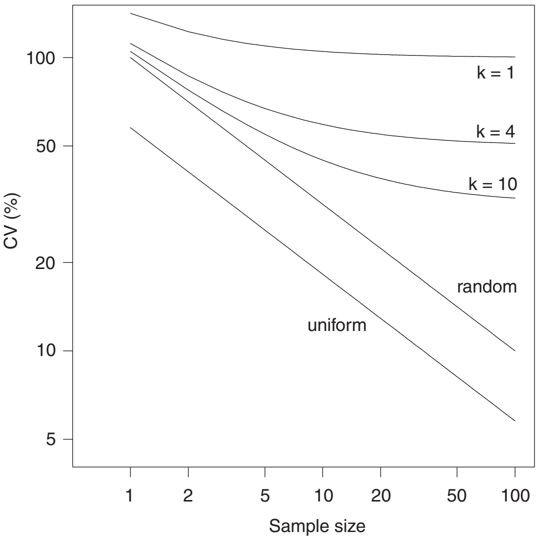

Fig. 1. Temporal variability in population samples resulting from spatial variance in temporally stable populations if individual are dispersed uniformly (spatial variance = 0.5 * mean abundance), randomly (spatial variance = mean abundance) or in an aggregated fashion (spatial variance > mean abundance). *k* refers to the parameter of negative binomial distribution (the lower the *k*, the higher the degree of aggregation).

Spatial variance in community samples is a more complicated problem. Most ecologists define and delimit communities based on taxonomic and spatial criteria: a community comprises populations of some taxonomically defined group of species coexisting at a site or in a region, i.e. within the sampling area. It may be that the spatial scale of the sampling unit we have chosen differently reflects the spatial distribution of species. Even if we defined community using taxonomically stringent criteria, e.g. song bird communities, our sampling area could encompass tens of home ranges (territories) of one species, but just a fraction of a home range of another species. Sherry and Holmes (1985), for example, showed that territory sizes may differ more than by an order of magnitude between bird species with approximately equal body size. The effects of spatial variance on the temporal variability vary accordingly.

## 4. Removing sampling variance

It is also perfectly possible that the data we have gathered is affected by both types of sampling variances. This combined variance is likely not additive, i.e. it is not simply a sum of two variance components, but there is an interaction between spatial variance and the variance due to inexactness of the counts. The spatial variance, we observe, is by necessity influenced by the sampling efficiency as e.g. completely uniformly dispersed population may appear randomly dispersed when $p_i < 1$. Likewise, spatial variance may affect sampling probability, e.g. $p_i$ being higher in hot-spots than in low density areas.

In most cases, therefore, we have no means to separate the effects of the two sources of the sampling variance, but their effects must be estimated simultaneously. There are sampling methods where either (very seldom both) source of the sampling variance is relatively small. Typical large-scale surveys such as the North American breeding bird survey (Robbins *et al.* 1986), the Finnish type line transect method for censusing and monitoring land birds (Järvinen & Väisänen 1983), the wildlife triangle scheme in Finland (Lindén *et al.* 1996), and the British butterfly monitoring scheme (Pollard & Yates 1993), encompass such large areas that the spatial variance in the populations of many species becomes averaged out, and all that remains is the inexactness of the counts. The territory mapping method (Anon. 1970) for estimating densities of breeding birds is considered very accurate (Nilsson 1977, Enemar *et al.* 1979), but comprehends spatially rather limited sampling units, and therefore temporal variability estimates are beset by spatial variance. A rule of thumb could be that if populations are spatially aggregated, as most natural populations obviously are, sampling units should be at least an order of scale larger than the scale of aggregation to avoid large bias. This, naturally, requires knowledge on the dispersion and its scale, which in turn means replication in space.

As discussed above, both sources of sampling variance generate patterns of decreasing temporal variability with the mean abundance. Low temporal variability in density tends to imply that density-dependent processes are operating (Hanski & Tiainen 1988, Hanski & Woiwod 1993). Empirical studies have shown that slopes in the Taylor power plots for terrestrial vertebrates, birds in particular, are smaller than for invertebrates (Hanski 1990). Low values of $b$, i.e. sharply decreasing variability with mean abundance, have been interpreted in terms of strong population regulation in birds (via territoriality, for example), but sampling error alone may produce such slopes. It would therefore be important to separate sampling variability from true variability.

If the counts ($X_{ti}$) have been replicated ($i = 1...n$) at each time period $t$ then sampling error can be estimated and removed (Link *et al.* 1994, McArdle & Gaston 1995, Stewart-Oaten *et al.* 1995). There are two possibilities: replication in space (more sampling plots) and in time (more samples per a plot). Link *et al.* (1994) compared the cost efficiency of replicating in time vs. in space, and concluded that replication in space is better in most situations (*see also* McArdle & Gaston 1995). Replication in space adds to the number of (independent) observations and still enables the sampling variance to be estimated.

After adequate replication, it is then in principle possible to use ANOVA techniques of variance component estimation to estimate the within-$t$ (the one that we want to remove) and among-$t$ (temporal variance that we want to estimate without error) variance. Observed variance then becomes:

$$\mathrm{var}(\log X_{ti}) = \mathrm{var}(\log N_t) + \mathrm{var}(\log X_{\cdot i}), \quad (2)$$

where the last term refers to within-$t$ variance of the counts in each replicate $i$. The true temporal variance, $\mathrm{var}(\log N_t)$, can then be solved. If any of the $X_{ti}$ values are zeros then some $\log(X_{ti})$'s remain undefined and the analysis must be performed on the untransformed data. The estimated $\mathrm{var}(N_t)$ can then be converted to a $\mathrm{CV}(N_t)$ (*see* McArdle & Gaston 1995).

Link *et al.* (1994) analyzed data on 98 species from the North American breeding bird surveys, and using repeated counts at survey sites removed the sampling variance. They showed that on average the sampling variability comprised 36% of the total variability ranging from 3.5 to 100%. For 14 species more than half of the variation in counts was attributable to sampling variance. The degree in error was clearly more in less abundant species.

Even if $X_t$'s are not replicated at each $t$ it is still possible to separate the sampling variance from 'true' temporal variance, but only if we are prepared to make assumptions on the distribution of $X_t$ (Järvinen & Lokki 1978, McArdle & Gaston

1995). If we assume that $X_t$ is distributed as a negative binomial then we can estimate the proportionate variation (Samways 1990), $CV^2(N_t)$, by the moment estimator:

$$CV^2(N_t) = CV^2(X_t) - 1/x, \qquad (3)$$

where $x$ is the mean of $X_t$. It should be noted that the negative binomial covers all cases where distribution is aggregated or random (Poisson distribution is a special case of a negative binomial with $c = 1/k = 0$). If we assumed that the distribution of $X_t$ is lognormal then we must estimate $CV(N_t)$ differently (*see* McArdle & Gaston 1995).

If a population is completely stable ($CV^2(N_t) = 0$), then all the variability is because of the sampling variance. From Eq. 2 we see that the observed variability $CV(X_t)$ then equals $\sqrt{}1/x_t$, which is the amount of error variability. As shown before, if the sampling variance is alone caused by inexactness of the counts, the variability that follows can be estimated as $\sqrt{}(1 - p_i)/p_iN_t$. It can clearly be seen that this variability term equals $\sqrt{}1/x_t$ if $p_i = 0.5$. This is not surprising since a negative binomial distribution can be thought of as a Poisson variable $X_t$, whose expected value is from gamma distribution (McArdle & Gaston 1995). If we were sure that all variability is because of the inexactness of the counts and that $p_i$ is very different from 0.5, then we obviously should use the term $\sqrt{}(1 - p_i)/p_iN_t$ to estimate error variability rather than the proportionate variation.

These methods are sensitive to assumption of the distribution of $X_t$ and to assumption of the sampling efficiency, $p_i$, respectively. If we want to test the appropriateness of the assumptions we would need to replicate sampling in time or in space. As already discussed above, in this case we have more direct and reliable methods of estimating sampling variance.

## 5. Examples

### 5.1. Does density vary more in northern bird communities than in southern ones?

Gaston and McArdle (1995; *see also* Stewart-Oaten *et al*. 1995) reviewed a number of patterns in the magnitude of temporal variability of populations and the evidence that exists for them. Because of the methodological drawbacks, most, if not all, empirical studies of such patterns are plagued with sampling errors and must, therefore, be interpreted with caution. We consider only some examples. We first deal with the classical question of a latitudinal variation in community variability, namely that northern communities are more variable than southern ones. Then we reanalyze existing data by removing sampling variance with a method just described. Finally, we discuss some recent publications on variability in core vs. peripheral populations and source vs. sink habitats, concluding that not enough caution has been exercised in interpreting results.

Northern animal populations are considered less stable than southern ones (MacArthur 1955). Likewise, variability in northern communities is considered greater than in southern ones, an idea which was originally derived from the 'diversity begets stability' principle (Elton 1958, Hutchinson 1959, *but see* May 1981). We used data on forest birds, gathered using two different methods: line transect (Järvinen & Väisänen 1983) and territory mapping (Anon. 1970). As discussed above, these two methods exemplify the two types of sampling variance.

As shown in Table 1, yearly sample sizes in the data sets varied between 96 and 744 pairs, and therefore, controlling for the effects of sampling variance is certainly required. As replicated samples are not available, we applied the proportional variation method to estimate temporal variability. Original variability values, $CV(X_t)$, show positive correlation with latitude ($r = 0.53$, $p = 0.03$). After removing the sampling variance, variability ($CV(N_t)$) was even more strongly correlated with latitude ($r = 0.58$, $p = 0.015$; Fig. 2). Note that in five data sets the sample size was too small to allow estimation of the temporal variability in pair numbers in these communities (negative values of $CV(N_t)$ were truncated zero in Fig. 2). This particularly applied to mapping data sets. Our results support Järvinen's (1979) results that, in the north, variability is greater than in the south.

Obviously latitude *per se* can have no direct effect on the temporal variability, but latitude itself is correlated with some other variables which have direct effects on variability. Järvinen (1979) concluded that environmental (climatic) unpredictability increases northwards and mainly causes

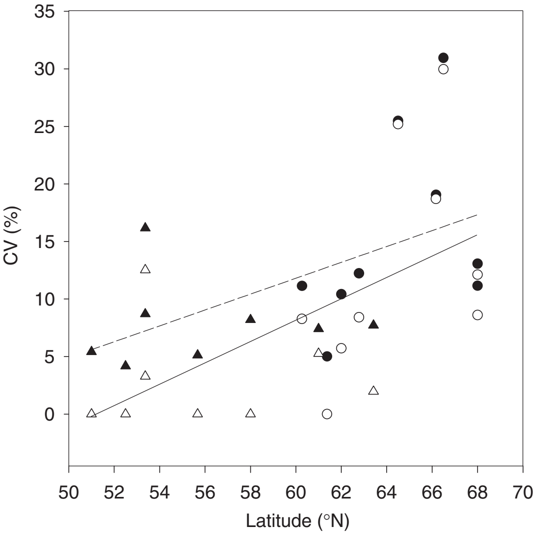

Fig. 2. Variability in total density of bird communities at different latitudes. Triangles refer to mapping data and circles to line transect data. Filled symbols denote observed CV($X_t$) values whereas open symbols represent variability (CV($N_t$)) after sampling variance was removed using proportionate variation method. Note that negative values of CV($N_t$) were truncated to zero.

the pattern. It is notable that Noon *et al*. (1985) found no support for northward increasing variability in North American bird communities. A possible explanation for this discrepancy is the different latitudinal range between studies. Järvinen (1979) particularly contrasted data from the temperate zone with those from the boreal. Also our data suggest that variability is high in the northernmost data sets, and relatively low and constant among sub-boreal bird communities (Fig. 2). Very little data in the Noon *et al*. (1985) study originated from the boreal zone.

## 5.2. Variability in central and peripheral populations

It has often been stated that populations at the edge of a species' range are less subject to density-dependent factors than those at the center, resulting in lower temporal variability in centrally located populations (e.g. Thomas *et al*. 1994). For example, Curnutt *et al*. (1996) suggested that populations of sparrows in North America varied more

Table 1. Data sets used in studying the correlation between density variability and latitude in forest bird communities. *N* refers to yearly average sample size (number of pairs observed) and Size to the area of the study plot (mapping data, in ha) or to the length of the transect (line transect data, km). Each data covers 4–6 consecutive years. If a study contained more than 6 yrs data, the latest 5 yrs data were selected.

| Source | Latitude, °N | *N* | Size |
|---|---|---|---|
| Mapping data | | | |
| Hogstad 1993 | 63°25´ | 180 | 100 |
| Solonen 1986 | 61°00´ | 367 | 30 |
| Morozov 1992 | 58°00´ | 108 | 19.5 |
| Enemar 1966 | 55°40´ | 108 | 13 |
| Tomialojc *et al*. 1984 | 53°22´ | 96 | 25 |
| Tomialojc *et al*. 1984 | 53°22´ | 154 | 24 |
| Witt 1974 | 52°30´ | 303 | 36.8 |
| Williamson & Williamson 1973 | 51°00´ | 317 | 58 |
| Line transect data | | | |
| Virkkala 1989 | 68°00´ | 199 | 8.1 |
| Virkkala 1989 | 68°00´ | 419 | 11.7 |
| Helle & Mönkkönen 1986 | 66°30´ | 165 | 11.1 |
| Mäkelä 1991 | 66°10´ | 744 | 25 |
| J. Inkeröinen & M. Mönkkönen, unpubl. | 64°30´ | 716 | 21 |
| E. Rajala, unpubl. | 62°47´ | 127 | 3.6 |
| E. Ylinen, unpubl. | 62°00´ | 132 | 3.3 |
| T. Kaasalainen, unpubl. | 61°23´ | 223 | 4.5 |
| A. Seppälä, unpubl. | 60°16´ | 181 | 5.1 |

in time when close to the edge of their ranges. This pattern may also be due to sampling error because mean abundances of edge populations were lower than those further away from the edge (*see also* Brown 1984). Curnutt *et al*. (1996) used North American breeding bird survey (BBS) data which applies a standardised single-visit method. In each survey, birds are censused on 50 3-min stops located 800 m apart along roads. As in all single-visit methods to census birds, sampling efficiency $p_i$ is less than 1. Curnutt *et al*. (1996) restricted their analyses to abundances higher than 1.5 birds per site to avoid sampling bias. This is a strictly arbitrary threshold, however, and as shown in this study mean sample sizes should be several tens of individuals if one wishes to avoid the extensive errors due to inefficient sampling (*see also* Link *et al*. 1994). No effort in Curnutt *et al*. (1996) was made to control for sampling variance above these threshold abundances. Moreover, Curnutt *et al*. (1996) replaced zero counts by 0.5. As shown by Anscombe (1948) and McArdle *et al*. (1990), the $\log(X_t + A)$ transformation inevitably biases variability estimates.

It may not be possible at all to gain very firm conclusions about variability differences between central and edge populations because so many variables may change along with the geographical positions, e.g. climate and other environmental conditions. Perhaps the best alternative is to study birth rates, individual survival and site tenacity in different populations; variability stems from these three factors.

### 5.3. Variability in sink and source populations

The sampling variance, if not taken into account, may cause serious problems in conservation biology as well. For example, in species conservation efforts, it is of crucial importance to know the areas or (sub)populations that are the most productive, serving as a source of surplus individuals to less productive areas/(sub)populations, called population sinks (Pulliam 1988). It has been suggested that one could identify sink habitats/populations from sources by higher temporal variability (e.g. Howe *et al*. 1991). The problem here is that population density in sink habitats is usually lower than in source patches. If sampling variance is ignored, this potentially leads to unsubstantiated conclusions and management that proves to be detrimental to the population; low variability may appear because of a large sample size alone.

Beshkarev *et al*. (1994) compared variability between hazel grouse (*Bonasia bonasa*) populations living in source and sink-dominated landscapes in NE European Russia, based on 30-yr. data. They found very little difference in temporal variability. They suggested that significant autocorrelation of density with a 1-yr. time lag would be a more powerful indicator of source habitats than temporal variability. The yearly samples of Beshkarev *et al*. (1994) in the sink-dominated landscape were rather small (8 individuals on average vs. 49 individuals in the source-dominated landscape), and coupled with $p_i < 1$ of the line transect method they used, it is not surprising that no autocorrelation was found in hazel grouse populations in the sinks; there is simply too much ‘noise’ in the samples.

Concluding, despite the extensive recent literature on problems involved in measuring variability in animal populations and communities, the message seems to a large extent to have been lost. It is increasingly clear that variability estimates are of little use if sampling variance is not removed and if thought is not given to how to interpret and compare variability estimates. Spatial modelling of sampling design could provide deeper insight into apparent and true temporal variability in animal populations and communities.

*Acknowledgements*: We thank the Academy of Finland for financial support (to MM), J. Kouki and two anonymous referees for fruitful comments on the manuscript. T. Kaasalainen, E. Rajala, A. Seppälä, and E. Ylinen kindly allowed us to use their unpublished data.

## References

Anon. 1970: Recommendations for an international standard for a mapping method in bird census work. — In: Svensson, S. (ed.), Bird census work and environmental monitoring: 49–52. Bull. Ecol. Res. Comm. no. 9, NFR, Stockholm.

Anscombe, F. J. 1948: The transformation of Poisson, binomial and negative binomial data. — Biometrika 35: 246–254.

Beshkarev, A. B., Swenson, J. E., Angelstam, P., Andrén, H. & Blagovidov, A. B. 1994: Long-term dynamics of hazel grouse populations in source- and sink-dominated pristine taiga landscapes. — Oikos 71: 375–380.

Brown, J. H. 1984: On the relationship between abundance and distribution of species. — Am. Nat. 124: 255–279.

Chesson, P. L. 1986: Environmental variation and the coexistence of species. — In: Diamond, J. & Case, T. J. (eds.), Community ecology: 240–256. Harper & Row, Publishers, New York.

Curnutt, J. L., Pimm, S. L. & Maurer, B. A. 1996: Population variability of sparrows in space and time. — Oikos 76: 131–144.

Elton, C. S. 1958: The ecology of invasion by animals and plants. — Methuen, London.

Enemar, A. 1966: A ten-year study on the size and composition of a breeding passerine bird community. — Vår Fågelvärld Suppl. 4: 47–94.

Enemar, A., Klaesson, P. & Sjöstrand, B. 1979: Accuracy and efficiency of mapping territorial willow warblers Phylloscopus trochilus. — Oikos 33: 176–181.

Gaston, K. J. 1988: Patterns in the local and regional dynamics of moth populations. — Oikos 53: 49–57.

Gaston, K. J. & McArdle, B. H. 1993: Measurement of variation in the size of populations in space and time: some points of clarification. — Oikos 68: 357–360.

Gaston, K. J. & McArdle, B. H. 1994: The temporal variability of animal abundances: measures, methods and patterns. — Phil. Trans. R. Soc. Lond. B 345: 335–358.

Hanski, I. 1982: On patterns of temporal and spatial variation in animal populations. — Ann. Zool. Fennici 19: 21–37.

Hanski, I. 1990: Density dependence, regulation and variability in animal populations. — Phil. Trans. R. Soc. Lond. B 330: 141–150.

Hanski, I. & Tiainen, J. 1988: Bird ecology and Taylor's variance-mean regression. — Ann. Zool. Fennici 26: 213–217.

Hanski, I. & Woiwod, I. P. 1993: Mean-related stochasticity and population variability. — Oikos 67: 29–39.

Hanski, I., Kouki, J. & Halkka, A. J. 1993: Three explanations of the positive relationship between distribution and abundance of species. — In: Ricklefs, R. E. & Schluter, D. (eds.), Species diversity in ecological communities: 108–116. Chicago University Press, Chicago.

Helle, P. & Mönkkönen, M. 1986: Annual fluctuations of land bird communities in different successional stages of boreal forest. — Ann. Zool. Fennici 23: 269–280.

Helle, P. & Pulliainen, E. 1983: On the efficiency of the line transect method: a study based on nest searching. — Ornis Fennica 60: 35–41.

Hogstad, O. 1993: Structure and dynamics of a passerine bird community in a spruce-dominated boreal forest. A 12-year study. — Ann. Zool. Fennici 30: 43–54.

Horne, J. K. & Schneider, D. C. 1995: Spatial variance in ecology. — Oikos 74: 18–26.

Howe, R. W., Davis, G. J. & Mosca, V. 1991: The demographic significance of sink populations. — Biol. Conserv. 57: 239–255.

Hutchinson, G. E. 1959: Homage to Santa Rosalia, or why are there so many kinds of animals. — Am. Nat. 95: 137–145.

Järvinen, O. 1979: Geographical gradients of stability in European land bird communities. — Oecologia 38: 51–69.

Järvinen, O. & Lokki, J. 1978: Indices of community structure in bird censuses based on a single visit: effect of variation in species efficiency. — Ornis Scand. 9: 87–93.

Järvinen, O. & Väisänen, R. A. 1983: Correction coefficients for line transect censuses of breeding birds. — Ornis Fennica 60: 97–104.

Leps, J. 1993: Taylor's power law and the measurement of variation in the size of populations in space and time. — Oikos 67: 349–356.

Lindén, H., Helle, E., Helle, P. & Wikman, M. 1996: Wildlife triangle scheme in Finland: methods and aims for monitoring wildlife populations. — Finnish Game Research 49: 4–11.

Link, W. A. & Nichols, J. D. 1994: On the importance of sampling variance to investigations of temporal variation in animal population size. — Oikos 69: 539–544.

Link, W. A., Barker, R. J., Sauer, J. R. & Droege, S. 1994: Within-site variability in surveys of wildlife populations. — Ecology 75: 1097–1108.

MacArthur, R. H. 1955: Fluctuating animal populations and a measure of community stability. — Ecology 36: 533–536.

Mäkelä, J. 1991: Kuusamon Valtavaaran seudun maalinnusto – linnuston rakenne ja vuosivaihtelu vuosina 1988–1992. — Metsähallituksen luonnonsuojelujulkaisuja, Sarja A: 20. [In Finnish with Swedish summary.]

May, R. M. 1981: Patterns in multi-species communities. — In: May, R. M. (ed.), Theoretical ecology: principles and applications, 2nd ed.: 197–227. Blackwell Scientific Publications, Oxford.

McArdle, B. H. & Gaston, K. J. 1992: Comparing population variabilities. — Oikos 64: 610–612.

McArdle, B. & Gaston, K. 1993: The temporal variability of populations. — Oikos 67: 187–191.

McArdle, B. & Gaston, K. J. 1995: The temporal variability of densities: back to basics. — Oikos 75: 165–171.

McArdle, B. H., Gaston, K. J. & Lawton, J. H. 1990: Variation in the size of animal populations: patterns, problems and artefacts. — J. Anim. Ecol. 59: 439–454.

Møller, A. P. 1992: Interspecific response to playback of bird song. — Ethology 90: 315–320.

Morozov, N. S. 1992: Breeding forest birds in the Valdai Uplands, north-west Russia: assemblage composition, interspecific associations and habitat amplitudes. — Ann. Zool. Fennici 29: 7–27.

Nilsson, S. G. 1977: Estimates of population density and changes for titmice, Nuthatch, and Treecreeper in southern Sweden – an evaluation of the territory mapping method. — Ornis Scand. 8: 9–16.

Noon, B. R., Dawson, D. K. & Kelly, J. P. 1985: A search for stability gradients in North American breeding bird communities. — Auk 102: 64–81.

Pielou, E. C. 1977: Mathematical ecology. — John Wiley & Sons, New York, USA.

Pimm, S. L. 1991: The balance of nature?: ecological issues in the conservation of species and communities. — Univ. Chicago Press, Chicago.

Pimm, S. L. 1993: Life on an intermittent edge. — Trends Ecol. Evol. 8: 45–46.

Pollard, E. & Yates, T. J. 1993: Monitoring butterflies for ecology and conservation. — Chapman and Hall, London.

Pulliam, H. R. 1988: Sources, sinks, and population regulation. — Am. Nat. 132: 652–661.

Pulliam, H. R. & Danielson, B. J. 1991: Sources, sinks, and habitat selection: a landscape perspective on population dynamics. — Am. Nat. 137: S50–S66.

Redfearn, A. & Pimm, S. L. 1988: Population variability and polyphagy in herbivorous insect communities. — Ecol. Monogr. 58: 39–55.

Robbins, C. S., Bystrak, D. & Geissler, P. H. 1986: The breeding bird survey: its first fifteen years, 1965–1979. — U.S. Fish & Wildlife Service, Resource Publ. 157.

Rosewell, J., Shorrocks, B. & Edwards, K. 1990: Competition on a divided and ephemeral resource: testing the assumptions. I. Aggregation. — J. Anim. Ecol. 59: 977–1001.

Samways, M. J. 1990: Species temporal variability: epigaeic ant assemblages and management for abundance and scarcity. — Oecologia 84: 482–490.

Schoener, T. W. 1986: Patterns in terrestrial vertebrate versus arthropod communities: do systematic differences in regularity exist? — In: Diamond, J. & Case, T. J. (eds.), Community ecology: 556–586. Harper & Row, Publishers, New York.

Sherry, T. W. & Holmes, R. T. 1985: Dispersion patterns and habitat responses of birds in northern hardwood forests. — In: Cody, M. L. (ed.), Habitat selection in birds: 283–309. Academic Press, New York.

Solonen, T. 1986: Stability of a bird community in southern Finland. — Ornis Scand. 17: 106–116.

Stewart-Oaten, A., Murdoch, W. W. & Walde, S. J. 1995: Estimation of temporal variability in populations. — Am. Nat. 146: 519–535.

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

Taylor, L. R. 1984: Assessing and interpreting the spatial distribution of insect populations. — Ann. Rev. Entomol. 29: 321–357.

Taylor, L. R. 1986: Synoptic dynamics, migration and the Rothamsted insect survey. — J. Anim. Ecol. 55: 1–38.

Taylor, L. R. & Woiwod, I. P. 1980: Temporal stability as a density dependent species characteristics. — J. Anim. Ecol. 49: 209–224.

Taylor, L. R. & Woiwod, I. P. 1982: Comparative synoptic dynamics. I. Relationships between inter- and intraspecific spatial and temporal variance/mean population parameters. — J. Anim. Ecol. 51: 879–906.

Taylor, L. R., Woivod, I. P. & Perry, J. N. 1978: The density dependence of spatial behaviour and the rarity of randomness. — J. Anim. Ecol. 47: 383–406.

Thomas, J. A., Moss, D. & Pollard, E. 1994: Increased fluctuations of butterfly populations towards the northern edges of species' ranges. — Ecography 17: 215–220.

Tomialojc, L., Wesolowski, T. & Walankiewicz, W. 1984: Breeding bird community of a primaeval temperate forest (Bialowieza National Park, Poland). — Acta Ornithol. 20: 241–310.

Verner, J. 1992: Data needs for avian conservation biology: have we avoided critical research? — Condor 94: 301–303.

Virkkala, R. 1989: Short-term fluctuations of bird communities and populations in virgin and managed forests in Northern Finland. — Ann. Zool. Fennici 26: 277–285.

Wiens, J. A. 1981: Scale problems in avian censusing. — Studies in Avian Biology 6: 513–621.

Williamson, R. & Williamson, K. 1973: The bird community of yew woodland at Kingley Vale, Sussex. — British Birds 66: 12–23.

Witt, K. 1974: 6-jährige Siedlungsdichteuntersuchung im Spandauer Forst/Berlin 1968–73. — In: Oelke, H. (ed.), 5. Deutsche Siedlungsdichte-Tagung: 73–78. Peine.

Xia, X. & Boonstra, R. 1992: Measuring temporal variability of population density: a critique. — Am. Nat. 140: 883–892.