

Quantifying temporal variability in population abundances

Joel P. Heath

Joel P. Heath (*jpheath@sfu.ca*), Centre for Wildlife Ecology, Dept Biological Sciences, Simon Fraser Univ., 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6.

Understanding variability of population abundances is of central concern to theoretical and applied evolutionary ecology, yet quantifying the conceptually simple idea has been substantially problematic. Standard statistical measures of variability are particularly biased by rare events, zero counts and other 'non-Gaussian' behaviour, which are often inappropriately weighted or excluded from analysis. I conjecture that these problems are primarily a function of calculating variation as deviation from an average abundance, while the average may not be static, nor actually reflect abundance at any point in the time series. Here I describe a simple metric (population variability PV) that quantifies variability as the average percent difference between all combinations of observed abundances. Zero counts can be included if desired. Similar to standard metrics, variability is measured on a proportional scale, facilitating comparative applications. Standard metrics are based on Gaussian distributions, are over-sensitive to rare events and heavy tailed behaviour, and can inappropriately indicate 'more time-more variation' effects (reddened spectrum). Here I demonstrate that, while PV behaves similarly for 'normal' time series, it is independent of deviation from mean abundance for heavy tailed distributions, its robustness to non-Gaussian behaviour resolves artificial reddened spectrum issues, and variability calculated using PV from short time series is substantially more accurate at estimating known long term variability than standard metrics. PV therefore provides common ground for evaluating the variability of populations undergoing different dynamics, and with different statistical distributions of abundance, and can be easily generalized to a variety of contexts and disciplines.

The importance of summarizing variability in population abundances is ubiquitous throughout evolutionary ecology, particularly in comparing population dynamics and evaluating extinction risk. There is general agreement such a metric be unit free and independent of the mean (i.e. measured on a proportional scale) to facilitate comparison (Pimm 1991, Gaston and McArdle 1994, McArdle and Gaston 1995, Inchausti and Halley 2002). As variance related to the mean (Taylor 1961), common techniques to calculate variability standardize logarithmically (standard deviation SD of the log transformed abundances; $SDL = SD[\log(N)]$), or arithmetically (coefficient of variation $CV = SD/\text{mean}$). Of course, normalizing for mean population size does not exclude variability related to density

dependent processes (this is often confused in the literature; Gaston and McArdle 1994).

SDL and CV are the most accepted variability metrics in ecology, being used for a variety of spatiotemporal comparisons. Although they measure variation on a proportional scale, they are not always independent of the mean (McArdle et al. 1990), and can be seriously biased by a number of issues, including zero counts, rare events and other 'non-Gaussian' behaviour of populations. Different characteristics and biases of SDL and CV suggest different scenarios may require different indices (Pimm 1991, McArdle and Gaston 1995), however this inhibits comparison, which is often the primary objective. It is therefore desirable a single metric be both applicable

and consistent across a wide range of scenarios. Many of these issues have been succinctly reviewed elsewhere (McArdle et al. 1990, Pimm 1991, Gaston and McArdle 1994).

Here I propose a technique for quantifying variability which is based on a conceptual approach of simply comparing all numbers in the time series. Statistical approaches first assume a particular form of distribution (e.g. 'normal'), calculate a measure of central tendency (e.g. mean) and subsequently calculate variation as average deviation from this 'central' value of population abundances (or the log transformation). While these statistical measures can also be calculated without using the time series mean (i.e. by using mean sum of squared differences), the proof of the equivalence of these approaches indicates that average deviation from the mean still underlies the metric, regardless of how it is calculated. I propose issues of non-Gaussian behaviour can be alleviated by simply comparing all numbers in the time series, and calculating average proportional difference. That is, instead of calculating variability as average deviation from average population size, variability of biological populations may be better described as average difference in abundance among years. I begin by discussing why standard statistical approaches are problematic for situations not meeting Gaussian assumptions. I then present the proposed method, 'population variability' PV describe several of its mathematical characteristics, and quantitatively illustrate that it behaves as desired: for 'normal' (Gaussian) time series, PV tightly corresponds to current statistical techniques (CV and SDL). However, instead of evaluating variability based on deviation from an arbitrary mean, PV more fairly treats both 'rare' and 'common' events by simply comparing all abundances relative to each other. Therefore, unlike standard statistical techniques, PV allows comparisons across a range of different dynamics and is robust to non-Gaussian behaviour, which I illustrate quantitatively using spectral analysis and simulation models.

Biological populations and Gaussian assumptions

Temporal variability has been defined as 'average deviation of a time series of abundances from an average value on a proportional scale' (Gaston and McArdle 1994). Both SDL and CV therefore assume that an average value of population size, and deviation from it, appropriately describe features of the population we are attempting to measure. While potentially suitable for some analysis, the validity of this assumption in reflecting stability/variability of biological populations is not clear. With the possible exception of populations driven by a stable equilibrium, deviation from an average abundance may not represent the

underlying dynamics which biologists are trying to capture when measuring variability. This could be particularly true if populations undergo cyclical dynamics or more complex behaviour. To properly facilitate comparisons among populations undergoing different dynamics, an index of variability should be robust to both 'normal' and non-Gaussian behaviour.

Even for relatively stable populations, deviation from average may not represent the variation ecologists are interested in. For example, consider a stable population that undergoes rare perturbations (e.g. a crash due to extreme environmental events, followed by quick rebound to carrying capacity via local production or emigration). In such scenarios, the mean will not reflect population size in ANY year, and variability calculated using SDL or CV will indicate deviation from the mean and therefore variability in ALL years (Fig. 1). As I will illustrate, if variability is measured by simply comparing abundance among all time steps (circles in Fig. 1), then like all events, rare events are naturally compared to every other event, rather than to an arbitrary Gaussian mean. This is intuitively appealing as rare events are by definition not Gaussian, but are expected to be important components of ecological systems, whereas most statistical approaches treat them as outliers and devise ways to exclude them. We can therefore consider 'rare' and 'common' as parts of a continuum of the underlying dynamics, rather than being discretely and arbitrarily defined (e.g. as a cutoff threshold for 'outliers', based on the inappropriate assumption of a Gaussian distribution).

As a further thought experiment, consider a stable population at a carrying capacity of 100 for 25 years

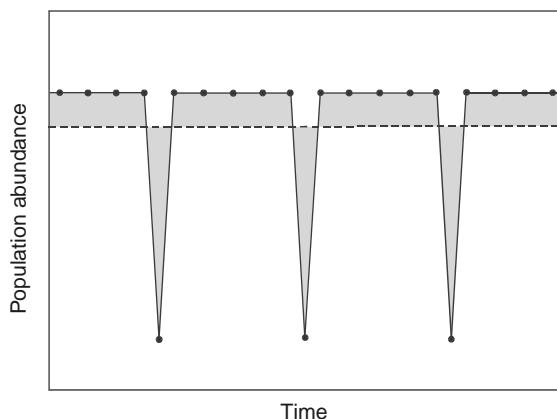


Fig. 1. For a population constant in abundance (circles) every year, except a few rare events, the mean (dashed line) does not reflect population size in any year, and variability quantified based on deviation (shaded area) implies variability in ALL years. Variability metrics that assume population abundances tend to fluctuate about a mean are therefore inappropriately sensitive to rare events.

that undergoes an sudden shift, due to some abrupt environmental change, but subsequently remains stable at a new carrying capacity of 10 for the next 50 years. Following this, it then returns to the old carrying capacity of 100 for 25 more years. A standard statistical approach would base variability on deviation in each year from a mean of 55 individuals, an abundance which is never achieved by the population, lies halfway between the bimodal distribution of abundances, therefore implies high variability in every year, and results in an extremely high value of variability ($CV = 0.82$; $SDL = 1.16$). In contrast, a comparison of all abundances indicates stability in 40% of comparisons, and more reasonable measure of variability, $PV = 0.45$. Of course, the same bimodal distribution of abundances and these same variability scores (PV, CV and SDL) would be obtained for a population oscillating between these two carrying capacities (100,10) every year (Note that, while the magnitude of difference between the two carrying capacities will determine the actual variability scores, in fact, exactly these same PV, CV, SDL scores will be obtained whenever the two carrying capacities differ by an order of magnitude). The chronology of events doesn't particularly matter for the type of variability we are discussing, we are only considering the events themselves; however, the time scale of censusing could be quite important in evaluating these two similar scenarios. Time series analysis would be necessary if the chronology of events are of interest.

To illustrate the effects of normal 'Gaussian' behaviour, rare events, a heavy-tailed (Cauchy) distribution, and a bimodal distribution of abundances on measures of variability Figure 2 summarizes 1000 randomizations of time series (100 intervals), for each of seven different simulations comparing CV, SDL and PV. For normal 'Gaussian' populations (abundances drawn randomly from a normal distribution of mean = 100; SD = 5 or 25), each variability metric similarly reflects the degree of stability over time. However, for stable normally distributed populations mean = 100; SD = 5) undergoing rare events (crash to mean = 10; SD = 5) at a frequency of 0.02, 0.05 and 0.1, CV and particularly SDL show exaggerated increases in variability. Similarly, for the heavy-tailed Cauchy distribution (limited to positive values, location parameter = 100, scale parameter = 5), CV and particularly SDL show extreme exaggerated variability. Finally, simulating the thought experiment of a changing carrying capacity/bimodal distribution of abundances (stable at mean 100; SD = 10 for 50 years and mean 10, SD = 1 for the remaining 50 years) indicates mean PV = 0.5 while CV and SDL inappropriately indicate extreme variability, even though the population was stable half the time. The biases inherent in standard statistical metrics (CV,SDL) are certainly clear. Following development of the PV technique, I will more formally

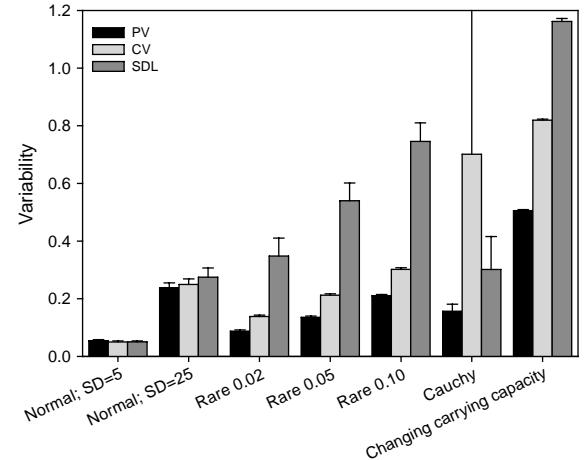


Fig. 2. Comparison of CV, SDL and PV across 7 simulations of population change for time series of 100 intervals. Each simulation was conducted 1000 times. For the first 2 simulations, 'normal' populations were generated by randomly drawing abundance from a normal distribution of mean = 100; standard deviation (SD) = 5 and 25, respectively. Rare events were simulated by selecting abundances from a stable distribution (mean = 100; SD = 5) but letting the population crash (mean = 10; SD = 5) at three frequencies (0.02, 0.05, 0.10), and by using the heavy tailed Cauchy distribution (location = 100, scale parameter = 5). The final simulation constructed bimodal time series of abundance, stable at mean = 100; SD = 10 for 50 years and mean 10, SD = 1 for the remaining 50 years, to illustrate the effects of an abrupt shift in carrying capacity or cyclical dynamics. Error bars are 95% confidence intervals. See text for further explanation.

compare these metrics and quantitatively illustrate that PV is more robust to different underlying dynamics and provides a more accurate measure of a biological concept of variability.

Calculation of population variability PV

Empirical time series of population abundances are usually discrete, a census being conducted at each time step. Sampling design and the grain (resolution) and extent (scope) of a data set determine the spatiotemporal scales of which it is representative. Measured abundances will include variability due to a variety of sources including sampling error, however this is an issue beyond the scope of this paper (e.g. Stewart-Oaten et al. 1995). To calculate variability in an intuitive manner, we can simply calculate the average proportional differences between all abundances in the time series. While comparing all pairs in a time series can be qualitatively redundant (because e.g. $a > b$ and $b > c$ tells us $a > c$) these comparisons are not quantitatively redundant. As we wish to quantitatively evaluate each

value of abundance relative to every other value of abundance, similar to metrics like Kendall's test for trend, we consider all possible combinations of abundance (C) which can be calculated from the number of time steps n in the time series as:

$$C = \frac{n(n - 1)!}{2} \quad (1)$$

We can then define z as the list of these possible pairwise comparisons; therefore $z = 1 \dots C$. Each z therefore represents a pair of time steps z_i and z_j to be compared by the difference function $D(z)$. This function simply calculates proportional difference between each z pair of abundances:

$$D(z) = \begin{cases} 0 & \text{if } z_i = z_j \\ \frac{\text{ABS}(z_i - z_j)}{\text{MAX}(z_i, z_j)} & \text{if } z_i \neq z_j \end{cases} \quad (2)$$

Using absolute value of the difference between z_i and z_j divided by the larger number simply ensures the proportion is the same even if we switch the values of z_i and z_j . That is, if $z_i = 10$ and $z_j = 100$, $D(z) = 90/100 = 90\%$, and likewise if $z_i = 100$ and $z_j = 10$, $D(z)$ is still $90/100 = 90\%$ (and not e.g. $90/10 = 900\%$). It is established that comparisons of absolute values, such as mean absolute deviation, provide more robust estimators (Press 1989). PV similarly uses absolute comparisons of abundance, however instead of standardizing using deviation from the mean, Eq. 2 standardizes by calculating the proportional difference between each value of abundance, i.e. the difference over the maximum. In fact, although Eq. 2 represents the framework in which PV was developed, it can be algebraically transformed to:

$$D(z) = 1 - \frac{\min(z_i, z_j)}{\max(z_i, z_j)} \quad (3)$$

indicating that it is based on a ratio comparison of each value in the time series. In this manner, abundance at every time step is compared with that at every other time step, yielding a distribution of proportional differences $D(z)$. While it may be interesting to investigate frequency distributions of $D(z)$ scores, most often an average will provide an adequate summary of population variability PV:

$$PV = \frac{\sum^z D(z)}{C} \quad (4)$$

We can therefore calculate variability based on a simple but thorough comparison of all abundances in a time series. Equation 2 calculates proportional difference, so the domain of $D(z)$ and PV is $[0, 1]$ in contrast with

SDL and CV, which, at least in theory, is $[0, \infty)$. A score of zero represents complete stability among years (i.e. $1/PV$ measures stability), while a value of 1 is approached as differences in population size approach infinity.

Like SDL and CV, the chronology of abundances is irrelevant and in fact, the PV approach quantifies differences across all time-lags. As chronology is irrelevant, autocorrelation structure will not influence the value of PV; the same time series could be randomly reorganized and lead to the same value of PV. However, it is noteworthy that sorting time series by abundance, rather than by time, can provide some insight into the behaviour of the different variability metrics. In this manner, trends in the ranked abundances can reflect underlying variability; for example, if sorted abundances exhibit no trend (constant abundance), then PV is equal to zero. In contrast, if the sorted abundances progress geometrically, then PV will approach 1. Intermediate between these extremes, PV will approach 0.5 when the sorted abundances approach an arithmetic progression, where the common difference of the progression and the lowest abundance have the same value. This can be understood by realizing that, in such cases, the frequency distribution of pair-wise proportional differences ($D(z)$ scores) will be evenly distributed. Again, note that because chronology is irrelevant, populations need not be exhibiting arithmetic or geometric growth to exhibit these patterns of variability; a multitude of chronologies including random fluctuation could produce the same pattern of sorted abundances. If prudent to the question, a time series could be de-trended before PV is quantified, although in many cases, trend might be considered an important component of variability. PV only indicates the overall stability/variability, and formal time series analysis should also be conducted if chronology of abundances is important to the question being investigated.

As dictated by the central limit theorem, PV values sampled repeatedly from a fixed distribution will be normally distributed. In order to quantify the influence of sampling error, repeated surveys would have to be conducted at each time step, to estimate confidence intervals for PV for a given population. Stewart-Oaten et al. (1995) discuss techniques to deal with sampling error, and further discussion is beyond the scope of this paper. The primary purpose of the PV technique is to allow comparisons of stability/variability among different populations. Given a sample of different populations, regression or other statistical analysis of PV and (for example) some ecological factor can be conducted in the standard manner, given the normal requirement that data/residuals meet the assumptions of the statistical test.

Ideally, there would be some gold standard against which to test the accuracy of PV. Lack of a 'true'

measure of variability precludes direct testing of any metric, making example analyses using real or population model data a relatively uninformative exercise. For this very reason, it is useful to have several different metrics in our toolbox which allow us to characterize different things about a time series, and a comparative approach can provide significant insight. As an important first step, it is desirable that PV behave similarly to CV and SDL for ‘well behaved’ Gaussian time series. PV, CV and SDL were calculated for 100 time series simulated with random mean [100,1000] and standard deviation [0,100] over 100 time steps. Correlation coefficients were calculated to evaluate concordance between PV, CV and SDL. This entire process was repeated 1000 times. PV was strongly correlated with both CV (mean $r = 0.997 \pm 0.001$ SD) and SDL (mean $r = 0.967 \pm 0.012$ SD). Therefore, PV behaves the same as CV and SDL for ‘normal’ (Gaussian) populations, as desired and expected. However, the necessity for an additional approach and a major advantage of PV is it’s robustness to non-Gaussian behaviour. I herein discuss how PV addresses and resolves several issues of non-Gaussian behaviour associated with CV and SDL.

The issue of zero counts

The presence of zero counts in biological time series is a major issue, and prevents using SDL as ($\log 0$) is undefined (Taylor 1961, McArdle et al. 1990). A first step is to consider the appropriateness of including zeros. Including data when the population is absent is uninformative if variability of the population is of interest but rather indicates variation of animals at a site (McArdle and Gaston 1993). In meta-population contexts, including local extinctions in sub-population variability could be important, particularly when evaluating local demographic differences (e.g. source-sink dynamics; Howe et al. 1991). When including zeros is important, often an arbitrary constant a is added (i.e. $SD[\log(N+a)]$), however this means variability is no longer on a proportional scale and is severely biased at low abundances (Gaston and McArdle 1994). PV calculates proportional differences using Eq. 2 which keeps $D(z)$ defined when either z_i or z_j equals zero. Such comparison between ‘extant’ and ‘extinct’ years is the only situation where $D(z) = 1$; otherwise $D(z)$ approaches 1 as the difference approaches infinity. The if condition ensures $D(z) = 0$ whenever $z_i = z_j$ including $z_i = z_j = 0$. If including zeros is important to the research question, then $z_i = z_j = 0$ can be biologically defined as “stably extinct”. It is important to note that I have been discussing true zeros: zeros due to sampling error will require careful consideration (McArdle and Gaston 1993, Gaston and McArdle 1994, McArdle and Gaston 1995).

Independent of average deviation from the mean

As mentioned in section 1, even though statistical measures of variability based on deviation from the mean can be calculated independent of the mean (i.e. by using the mean sum of squared differences), the proof of the equivalence of these approaches indicates that average deviation from the mean still underlies these metrics, regardless of how they are calculated. While no such proof exists for PV, simulations can be used to determine if PV is correlated with average deviation from the mean (ADM). Of course, given it has been demonstrated that PV behaves similar to CV and SDL for ‘normal’ populations, we would expect it to be correlated to ADM for Gaussian data, even though, unlike SDL and CV, it is not fundamentally based on deviation from the mean. However, for very non-Gaussian data, such as the heavy tailed Cauchy distribution, PV should not be correlated with ADM unless ADM somehow some how underlies PV. (Note that, although the mean and ADM can be calculated for any set of numbers/abundances, the mean is actually undefined for the Cauchy distribution). 100 time series of 100 units were generated to fit the normal distribution (mean [1000–4000], SD [10–90]) and the heavy-tailed Cauchy distribution) location [1000, 5000], scalar [10, 100]). For each time series, PV, CV and SDL were calculated, and correlation with ADM was determined across the 100 time series. This entire process was repeated 2000 times for each distribution. Figure 3 indicates, that as expected, each metric is correlated with average deviation from the mean for the normal distribution. However, unlike SDL and particularly CV, PV is not correlated with average deviation from the mean for the very non-Gaussian Cauchy distribution. While indirect, and not as elegant as a mathematical proof, this result indicates that average deviation from the mean is not an underlying feature of PV. As discussed previously, this is desirable for considering ‘rare’ events and comparing populations undergoing different dynamics, including those exhibiting non-Gaussian patterns of abundance.

Rare events and ‘more time–more variation’

Extreme sensitivity of SDL and CV to rare events has been recognized as a serious issue (for example, see Pimm’s (1991) analysis of song thrush populations which crashed following an extreme winter). While it may be tempting to exclude such ‘outliers’, this is undesirable as both rare and common events are expected to be important in ecology (Halley 1996). The remainder of this section and the next deal with devising a quantitative way to illustrate that PV is a more ‘robust’ measure than CV or SDL, in that it

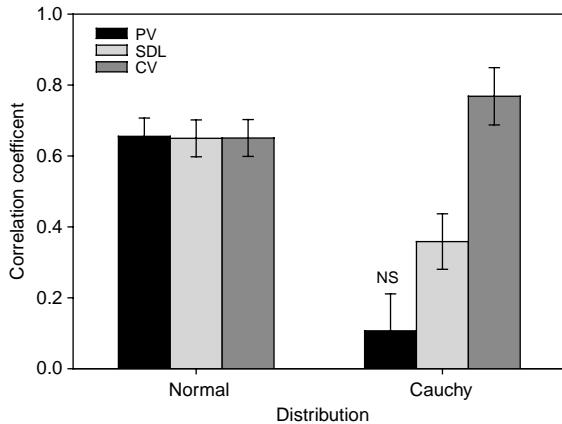


Fig. 3. Correlation of PV, CV and SLD with average deviation from the mean (ADM), across 2000 simulations of 100 time series, for each of the normal distribution (mean [1000,5000]; SD [10,90]), and the heavy tailed Cauchy distribution (location [1000,5000]; scale parameter [10,100]). While each metric is correlated with ADM for the normal distribution, PV is not correlated with ADM for the very non-Gaussian Cauchy distribution, indicating that average deviation from the mean does not underlie calculation of PV. Error bars are standard deviations.

functions more appropriately across time scales, and is not over-sensitive to rare events, but rather handles them in a fair and desirable manner.

An exciting recent line of research is focussed on investigating coupling between environmental and biological variation with important implications for understanding responses to environmental change. Many physical processes show scaling relationships with time, exhibiting more variation as time scale increases (Steele 1985; this has been termed spectral ‘reddening’, with analogy to visible light, to reflect the dominant importance of low frequency processes; Schneider 1994, Halley 1996). It has therefore been of interest to ask if reddened physical processes can force more time-more variation effects in ecological systems, in contrast to common assumption that ecological stochasticity follows a white-noise model (equal contribution across frequencies).

More time-more variation (reddening) of population abundances has been observed across a wide variety of taxa (Pimm and Redfearn 1988, Arino and Pimm 1995, Cyr 1997, Inchausti and Halley 2002) often using SLD or CV to measure variability. However, rare events can produce reddening in otherwise stationary time series (Mandelbrot 1999, Inchausti and Halley 2002). Spectral analysis more accurately evaluates reddening and therefore provides a gold standard against which to compare and evaluate the robustness of SLD, CV and PV. Therefore, time series of stable populations undergoing rare events were generated for 100 time

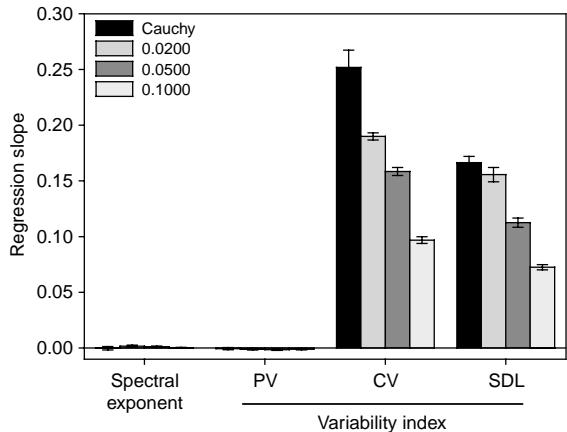


Fig. 4. Comparison of variance exponents (regression slope of log variability versus log time scale) among PV, CV, and SLD for simulated populations undergoing rare events at frequencies of 0.02, 0.05 and 0.1, and the heavy tailed Cauchy distribution (location [1000,5000], scale parameter [10,100]). Although all time series fit a white-noise model (spectral exponent ≈ 0), CV and SLD inappropriately indicate more time-more variation due to over-sensitivity to rare events (the lower the frequency of rare events, the more pronounced the effect). In contrast, PV is robust (not overly sensitive) to rare events, and is consistent with the gold standard of spectral analysis. Error bars are 95% confidence intervals. See text for details.

steps. Abundance at each time step was randomly selected from normal distributions of known mean and standard deviation. Rare-crashes (stable at 100 ± 5 , crashed to 10 ± 5) were simulated for 1000 time series, and rare-outbursts (stable at 10 ± 5 , outburst to 100 ± 5) for an additional 1000 time series. Rare events were set to occur at a frequency of 0.02, 0.05 and 0.1. An additional 2000 time series were simulated to fit the heavy tailed Cauchy distribution (random location parameter [1000,5000]; random scale parameter [10,100]). For each randomization (8000 time series), variability was measured over all time scales from 3 to 100 time steps using variance growth exponents after Inchausti and Halley (2002), with each of SLD, CV and PV as metrics. Variance exponents γ measure increase in variability over time quantified as regression slope of log variability against log time scale $\gamma = 0$ indicates white noise; $0 < \gamma < 1$ indicates decelerated increasing variance; $\gamma = 1$ indicates a random walk; $\gamma > 1$ indicates accelerating increasing variance; see Inchausti and Halley 2002 for details). Spectral exponents (SE; zero indicates white-noise, greater values indicate reddening) were calculated as (negative) regression slope of log spectral density versus log frequency from spectral analysis (Inchausti and Halley 2002). SE therefore provided a standard to confirm simulated

time series met white noise criteria ($SE < |\pm 0.1|$), and against which to evaluate and compare CV, SDL and PV.

Figure 4 summarizes results for each frequency of rare events and the Cauchy distribution, for each variability metric, clearly indicating that in all cases both SDL and CV inappropriately show strong increases in variation with time scale (more time—more variation) for these stationary time series (see also Mandelbrot 1999, Inchausti and Halley 2002; note that while Fig. 3 indicates increasing apparent reddening with decreasing frequency of rare events for CV and SDL, this is already established and not the purpose of this analysis). In contrast, PV correctly indicates constant variability across time scales, as confirmed by spectral exponents. Note that, although these results indicate PV is more appropriate than CV or SDL for evaluating more time-more variation, spectral analysis is the preferred method; the important result is that, unlike CV and SDL, PV appropriately handles rare events and is robust to non-Gaussian behaviour.

It is also important to ensure that PV behaves appropriately for detecting true instances of more time-more variation. Simulations were therefore conducted for 5000 random 100 year time series generated to fit a reddened noise model (using Fourier filtering to produce fractional Brownian motion of a given spectral density; after Crownover 1995). Spectral exponents SE were again used to confirm the true spectral density of each simulated time series ($SE > 0.2$; mean $SE = 0.89 \pm 0.23$ SD) and PV, CV and SDL were used to estimate more time-more variation, following the method of Inchausti and Halley (2002), as above. Regression slopes using PV, CV and SDL were highly correlated ($r > 0.97$), and like CV and SDL, PV correctly indicated true more time-more variation, and was highly correlated to the spectral exponent SE ($r = 0.71$, $p < 0.0001$). Therefore, while all variability metrics indicate true more time-more variation, because of over-sensitivity to rare events, CV and SDL inappropriately indicate more time-more variation for some stationary populations. In contrast, PV was consistent with spectral analysis in each scenario indicating its robustness to non-Gaussian behaviour.

Short-term estimates of long-term variability

A common problem in empirical ecology is that we are often required to estimate long term patterns based on short term data sets. Although some populations may exhibit more time-more variation effects, it is interesting to explore how variability measured in the short term reflects long term variability of a known time series. Given the previous analysis indicated that PV performs well across time scales, it seemed possible that

PV could provide more accurate estimates of long term variability from short term data sets. Therefore, for each of the normal (mean [100,1000], SD [10,1000]), lognormal (letter mu [0, 2], letter sigma [1, 2]) and Cauchy (location [1000, 5000], scale parameter [10, 100]) distributions, I simulated a 100 year time series, then randomly sub-sampled 100 times each for sample sizes from 3 to 99. Each metric (PV, CV, SDL) was calculated for the known (100 year) time series, and for each sub-sample (and averaged across sub-sampling randomizations). For each sub-sampling interval (3 to 99), the percent difference between the estimate and known long term variability was calculated. This entire process was repeated 1000 times for each distribution. The results are presented in Fig. 5 and clearly indicate that PV is substantially more accurate at estimating long term variability from small samples in each circumstance. For the normal distribution (Fig. 5a), PV is notably more accurate than SDL, whereas for the lognormal distribution (Fig. 5b), PV is much more accurate than CV (almost 40% for small samples) and similarly (though to a lesser magnitude) more accurate than SDL. Both CV and SDL are poor at estimating long term variability for the Cauchy distribution (Fig. 5c), their accuracy only linearly increasing with time scale, while the accuracy of PV accelerates with increasing time scale (note the difference in the scale of the y-axes). This clearly demonstrates that PV is not only robust to non-Gaussian behaviour, but further allows more accurate estimation of long term variability from shorter term data sets. It therefore also allows more robust comparisons of variability across time series/data sets of different lengths, though this should be done in a cautionary manner, particularly given the potential for (true) more time-more variation effects.

Conclusions

The variability of population abundances is recognized as one of the most important conceptual ideas in theoretical and empirical ecology, particularly for evaluating extinction risk and identifying, managing and developing conservation strategies for demographically and genetically important populations. Unfortunately the concept is more complex than often realised (Pimm 1991, Gaston and McArdle 1994), putting many published analyses in question (McArdle et al. 1990). Biological systems can exhibit a wide diversity of behaviour for which standard statistical approaches may not always be appropriate. The analysis presented in here certainly indicates PV has many appealing properties and advantages over existing metrics, justifying further exploration. Of course, given their underlying assumptions are met, statistical approaches and their

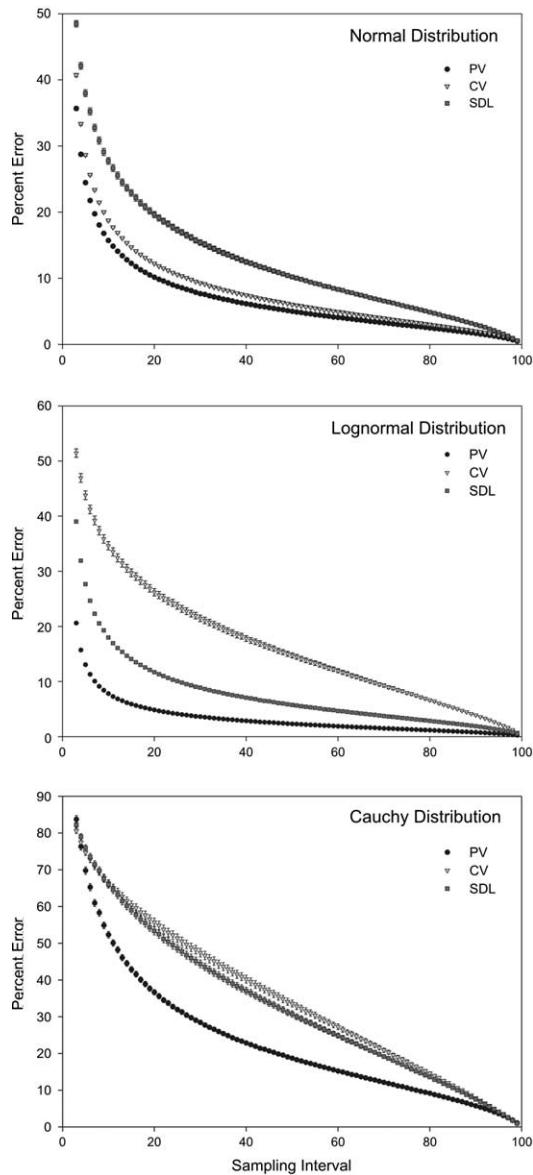


Fig. 5. A comparison of PV, CV and SDL at estimating long term known variability from shorter term sub-sampled data sets, for 1000 time series for each of the normal, lognormal and Cauchy distributions. For each distribution, PV was substantially more accurate (lower percent error, across sampling intervals) at estimating known long term variability than CV or SDL. SDL had particularly high error rates for the normal distribution, while CV was substantially inaccurate for the lognormal distribution. Both CV and SDL performed poorly for the Cauchy distribution, while the accuracy of PV accelerated with sampling interval for each distribution. Error bars are 95% confidence intervals, see text for details.

well defined distributional properties will still be important for addressing many ecological questions. PV is not being proposed as a mutually exclusive approach, and will likely be most informative when

used in conjunction with other statistical and time series techniques. Likewise, metrics like SDL and CV will no doubt still have a place in ecological analyses; however, for situations where consideration of rare events and other non-Gaussian behaviour are important, or where long term variability needs to be estimated from short term data sets, PV provides a more appropriate and robust metric. In addition to its conceptually simple and intuitive approach of comparing all abundances, it is compelling that PV is robust to a wide variety of both 'normal' and non-Gaussian behaviour. This provides common ground for comparing populations undergoing different dynamics in a conceptual framework which could also be useful for measuring stability/variability in a wide range of other contexts and disciplines.

Acknowledgements – Helpful comments at various stages of development were provided by (alphabetically) Wendell Challenger, Brian McArdle, Richard Lockhart, Bill Montevecchi, Patrick Nosil, Erikur Palsson, Greg Robertson, Bernie Roitberg, Dave Schneider and Carl Schwarz. Funding was provided by a Memorial Univ. of Newfoundland graduate fellowship and NSERC PGS-B Scholarship at Simon Fraser Univ.

References

- Arino, A. and Pimm, S. L. 1995. On the nature of population extremes. – *Evol. Ecol.* 9: 429–443.
- Crownover, R. M. 1995. Introduction to fractals and chaos. – Jones and Bartlett.
- Cyr, H. 1997. Does inter-annual variability in population density increase with time? – *Oikos* 79: 549–558.
- Gaston, K. J. and McArdle, B. H. 1994. The temporal variability of animal abundances: measures, methods and patterns. – *Philos. Trans. R. Soc. Lond. B.* 345: 335–358.
- Halley, J., M. 1996. Ecology, evolution and 1/f-noise. – *Trends Ecol. Evol.* 11: 33–37.
- Howe, R. W., Davis, G. J. and Mosca, V. 1991. The demographic significance of 'sink' populations. – *Biol. Conserv.* 57: 239–255.
- Inchausti, P. and Halley, J. 2002. The long-term temporal variability and spectral colour of animal populations. – *Evol. Ecol. Res.* 4: 1033–1048.
- Mandelbrot, B. B. 1999. Multifractals and 1/f noise. – Springer.
- McArdle, B. H. and Gaston, K.J. 1993. The temporal variability of populations. – *Oikos* 67: 187–191.
- McArdle, B. H and Gaston, K. J. 1995. The temporal variability of populations: back to basics. – *Oikos* 74: 165–171.
- McArdle, B. H., Gaston, K. J. and Lawton, J. H. 1990. Variation in the size of animal populations: patterns, problems and artefacts. – *J. Anim. Ecol.* 59: 439–454.
- Pimm, S. L. 1991. The balance of nature?. – Univ. Chicago Press.

Pimm, S. L. and Redfearn, A. 1988. Population variability of population densities. – *Nature* 334: 613–614.

Press, W. H., Flannery, B. P., Teukolsky, S. A. et al. 1989. Numerical recipes: the art of scientific computing. – Cambridge Univ. Press.

Schneider, D. C. 1994. Quantitative ecology: spatial and temporal scaling. – Academic Press.

Steele, J. H. 1985. A comparison of terrestrial and marine ecological systems. – *Nature* 313: 355–358.

Stewart-Oaten, A., Murdoch, W. W. and 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.

Appendix 1. Matlab equation for automatic calculation of the proposed metrics.

% calculates population variability, PV, from a vector
% of population abundances, P

```
C = nchoosek(P,2);
y = length(P);
Z = nchoosek(y,2);
for m = 1:Z
Num(m) = abs((C(m,1))-(C(m,2)));
Denom(m) = max((C(m,1)),(C(m,2)));
Diff(m) = Num(m)/Denom(m);
end
PV = mean(Diff)
clear C; clear Denom; clear Num; clear Z; clear i;
clear y; clear Diff;
```