On the use of auxiliary information on measurement error to improve the estimation of environmental variance in a Gompertz time series.

Matthew Ferrari

9 May 2002

1 Introduction

The use of mathematical models to describe population phenomena has been central to the fields of demographics and population biology. A perennial question in the description of biological populations is whether or not populations are self-limiting, or density dependent. In its broadest sense, density dependence refers to phenomena in which per capita growth rate is a non-constant function of population size. Most commonly, density dependent growth refers to the case in which per capita growth rate is a decreasing function of population size (though models of growth rate as an increasing function of population size (Allee effect) are used for small populations [6]). Though difficult to prove empirically, density dependent demographic phenomena have a certain logical appeal [16, 2]. When the number of individuals is low, there are more available resources per capita and the population will tend to increase. As the number of individuals increases, there are fewer available resources per individual which should result in lower fecundity and higher mortality. When the population becomes large enough, the per capita birth rate equals the per capita death rate and the population is maintained at a stable equilibrium referred to as carrying capacity. Models of density dependence are also attractive because of their flexibility, even in deterministic formulations, to describe a broad spectrum of qualitative phenomena observed in real time series; logistic growth to a stable equilibrium, damped oscillations, stable limit cycles, and chaos [1]. Despite its intuitive appeal, the role played by density dependent demographic phenomena in the regulation and persistence of biological populations has been debated in the ecological literature and many methods have been proposed to test for and estimate the magnitude of density dependence [7, 12, 18].

While deterministic models are useful for describing the qualitative behavior of biological time series, real time series exhibit stochastic behavior. The sources of variation in population numbers can be broadly categorized as environmental stochasticity or demographic stochasticity [8]. Environmental stochasticity refers to the population response to variation in weather patterns and resource availability that tend to affect the population as a whole. Populations are composed of individuals for whom reproduction and survival cannot be predicted. For small populations, this will result in variation in population growth rate. As populations become large, the mean per capita contribution to the population growth rate will have variance close to zero. This aspect of stochasticity, which tends to diminish as the population increases in number, is referred to as demographic stochasticity [8].

A third source of variability introduced in observed population time series is measurement error. In all but the most trivial cases, population time series are composed of data on estimated population sizes. Even cases that are purported to be full censuses of the population are often only approximations of total population size. If the measurement process is unbiased, the observed population size at time t, Y_t, can be thought of as a random variable with expected value equal to the true population size, N_t, and variance τ^2 (hereafter called measurement variance). The total variance in a population time series can be decomposed as some function of the variance associated with the population dynamic processes (environmental and demographic stochasticity), σ^2 (hereafter called process variance), and measurement variance. The value of this function will necessarily be greater than or equal to the process variance. Thus, the observed variance in a time series measured with error will tend to overstate the variance in the population demographic process [3].

The persistence of biological populations through time results from the interaction of deterministic processes and stochastic perturbations. A population with a positive growth rate can be driven to extinction if it is highly responsive to stochastic fluctuation [17, 11, 14]. The strength and form of density dependent phenomena can also influence the risk of extinction [9]. It is often the goal of conservation and wildlife management to minimize the risk of future extinction. Modeling and forecasting future population numbers form an essential part of such management [3]. To do so requires the use of empirical data to discern the form of the population dynamics and to estimate the parameters of the growth function. If the only data available are the observed population sizes, Y_t, then the process variance and measurement variance will be confounded and possibly unidentifiable. If the measurement variance is not accounted for, estimators of the process variance based on the observed variance in the time series will be positively biased. This presents distinct problems in the estimation of the process variance, and may affect the form of the population dynamic model chosen and the estimation of growth rate parameters.

If the measurement error can be estimated from some independent, auxiliary data, $\mathbf{U}$, then this estimate can be used to decompose the total variance in the observed time series into process and measurement components. One approach, pseudo likelihood, would be to estimate the measurement variance parameter, τ^2, from the auxiliary data, and then treat that estimate as known in the likelihood function for the observed time series and maximize the likelihood for the remaining parameters. Under general conditions, such pseudo maximum likelihood estimates are consistent and asymptotically normal [10]. Pseudo maximum likelihood has been criticized; however, because no compensation is made for the loss of degrees of freedom associated with the nuisance parameter, τ^2 [5]. If the auxiliary data can be expressed in a likelihood, and the data are independent of the observed time series, then the product of the two likelihoods will yield a joint likelihood for the time series and the auxiliary data which will account for the additional uncertainty in the estimate of τ^2.

In this paper I first illustrate the impact of auxiliary data on the width of the confidence intervals for additive variance components in a zero mean, normal process. I then examine the effect of incorporating auxiliary data, via pseudo likelihood and using a full likelihood, on the point and interval estimates of the process variance for a density dependent population growth model (Gompertz model). Lastly, I consider the effect of measurement error on the maximum likelihood estimates of the growth rate parameters of the Gompertz population growth model.

2 Estimation of confounded variance components for a zero mean normal process.

2.1 Point estimation of variance components

Consider a zero mean random phenomenon, X, with normally distributed deviations; X $\sim N(0, \sigma^2 + \tau^2)$. If $\mathbf{x} = (x_1, x_2, \ldots, x_{n_1})'$ is a random sample of size n_1 observations of this phenomenon, unique maximum likelihood estimates of the parameters σ^2 and τ^2 do not exist. Let U be an independent random variable with distribution, U $\sim N(0, \tau^2)$, and $\mathbf{u} = (u_1, u_2, \ldots, u_{n_2})'$ be a random sample of size n_2 from that distribution, and $\mathbf{u}$ be independent of $\mathbf{x}$.

The joint likelihood function for $\mathbf{x}$ and $\mathbf{u}$ given σ^2 and τ^2 is:

$$
\begin{aligned}
L(\sigma^2, \tau^2 \mid \mathbf{x}, \mathbf{u}) &= L(\sigma^2, \tau^2 \mid \mathbf{x}) L(\tau^2 \mid \mathbf{u}) \\
&= \frac{\exp\left[\frac{-\sum x_i^2}{2(\sigma^2+\tau^2)} - \frac{\sum u_j^2}{2(\tau^2)}\right]}{(\sigma^2+\tau^2)^{\frac{n_1}{2}} (\tau^2)^{\frac{n_2}{2}}} \\
\ln L(\sigma^2, \tau^2 \mid \mathbf{x}, \mathbf{u}) &= -\frac{1}{2(\sigma^2+\tau^2)} \sum_{i=1}^{n_1} x_i^2 - \frac{1}{2(\sigma^2)} \sum_{j=1}^{n_2} u_j^2 - \frac{n_1}{2} \ln(\sigma^2+\tau^2) - \frac{n_2}{2} \ln(\tau^2) \qquad (1)
\end{aligned}
$$

The maximum likelihood estimators of the parameters σ^2 and τ^2 are given below. Note that without the use of the auxiliary data only the quantity $\theta^2 = \sigma^2 + \tau^2$ can be estimated.

$$
\hat{\theta}^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i^2 \qquad (2)
$$

$$\hat{\tau^2} = \frac{1}{n_2}\sum_{j=1}^{n_2} u_j^2 \tag{3}$$

$$\hat{\sigma^2} = \hat{\theta^2} - \hat{\tau^2} \tag{4}$$

$$= \frac{1}{n_1}\sum_{i=1}^{n_1} x_i^2 - \frac{1}{n_2}\sum_{j=1}^{n_2} u_j^2 \tag{5}$$

Let ψ be the parameter vector $(\tau^2, \sigma^2)'$. Then the information matrix is the expected matrix of minus the second partial derivatives of the log likelihood function :

$$\mathbf{I}(\psi; \mathbf{x}, \mathbf{u}) = E\begin{bmatrix} \frac{\sum_{i=1}^{n_1} x_i^2}{(\sigma^2+\tau^2)^3} + \frac{\sum_{j=1}^{n_2} u_j^2}{(\tau^2)^3} - \frac{n_1}{2(\sigma^2+\tau^2)^2} - \frac{n_2}{2(\tau^2)^2} & \frac{\sum_{i=1}^{n_1} x_i^2}{(\sigma^2+\tau^2)^3} - \frac{n_1}{2(\sigma^2+\tau^2)^2} \\ \frac{\sum_{i=1}^{n_1} x_i^2}{(\sigma^2+\tau^2)^3} - \frac{n_1}{2(\sigma^2+\tau^2)^2} & \frac{\sum_{i=1}^{n_1} x_i^2}{(\sigma^2+\tau^2)^3} - \frac{n_1}{2(\sigma^2+\tau^2)^2} \end{bmatrix}. \tag{6}$$

The information matrix, evaluated at the maximum likelihood estimates is the observed information matrix,

$$\mathbf{I}(\hat{\psi}; \mathbf{x}, \mathbf{u}) = \begin{bmatrix} \frac{n_1^3}{2(\sum_{i=1}^{n_1} x_i^2)^2} + \frac{n_2^3}{2(\sum_{j=1}^{n_2} u_j^2)^2} & \frac{n_1^3}{2(\sum_{i=1}^{n_1} x_i^2)^2} \\ \frac{n_1^3}{2(\sum_{i=1}^{n_1} x_i^2)^2} & \frac{n_1^3}{2(\sum_{i=1}^{n_1} x_i^2)^2} \end{bmatrix}, \tag{7}$$

which can be rewritten as,

$$\mathbf{I}(\hat{\psi}; \mathbf{x}, \mathbf{y}) = \begin{bmatrix} a+b & a \\ a & a \end{bmatrix},$$

where

$$a = \frac{n_1}{(2(\hat{\tau^2} + \hat{\sigma^2}))^2}$$

$$b = \frac{n_2}{(2(\hat{\tau^2}))^2}.$$

Under general conditions, maximum likelihood estimates are asymptotically efficient with covariance matrix that approaches $\mathbf{I}^{-1}(\psi)$ [15] . Thus the inverse of the observed information matrix, evaluated at the maximum likelihood estimates, can be used to estimate the covariance matrix of the maximum likelihood estimators. Thus the variance of $\hat{\psi}$ can be approximated by

$$\mathbf{I}^{-1}(\hat{\psi}) = \frac{1}{ab}\begin{pmatrix} a & -a \\ -a & a+b \end{pmatrix}$$

$$Var(\hat{\psi}) \approx \mathbf{I}^{-1}(\hat{\psi}) = \begin{bmatrix} \frac{2(\hat{\tau^2})^2}{n_2} & -\frac{2(\hat{\tau^2})^2}{n_2} \\ -\frac{2(\hat{\tau^2})^2}{n_2} & \frac{2(\hat{\tau^2})^2}{n_2} + \frac{2(\hat{\sigma^2}+\hat{\tau^2})^2}{n_1} \end{bmatrix}. \quad (8)$$

Inspection of the information matrix shows that the variance term associated with the estimate of σ^2 is composed of two elements, one from the data vector $\mathbf{x}$ and one from the data vector $\mathbf{u}$. As n_2 increases, increasing the information about τ^2, the variance term associated with σ^2 decreases in magnitude and converges to $\frac{2(\hat{\sigma}^2+\hat{\tau}^2)}{n_1}$ as the sample size n_2 goes to infinity. If τ^2 is very small, the increased sampling of $\mathbf{u}$ will result in only small reductions in the standard error of $\hat{\sigma}^2$ (at most the reduction will be $2(\hat{\tau}^2)^2$). If τ^2 is large, the addition of information on $\mathbf{u}$ can substantially reduce the standard error of $\hat{\sigma}^2$. However, $\hat{\tau}^2$ remains as a term in the standard error of $\hat{\sigma}^2$ even when $n_2 = \infty$, and larger values of $\hat{\tau}^2$ will result in a larger standard error of $\hat{\sigma}^2$ for constant n_1.

Using the above estimated standard error for $\hat{\sigma^2}$, an approximate $(1-\alpha)100\%$ confidence interval for σ^2 for large n_1 and n_2 is,

$$\hat{\sigma^2} + Z_{\alpha/2}\sqrt{\frac{2(\hat{\tau^2})^2}{n_2} + \frac{2(\hat{\sigma^2}+\hat{\tau^2})^2}{n_1}} \le \sigma^2 \le \hat{\sigma^2} + Z_{1-\alpha/2}\sqrt{\frac{2(\hat{\tau^2})^2}{n_2} + \frac{2(\hat{\sigma^2}+\hat{\tau^2})^2}{n_1}}, \quad (9)$$

where $Z_{\alpha/2}$ is the $\alpha/2^{th}$ percentile of a standard normal distribution.

The interval given in equation 9 is a function of the sample sizes n_1 and n_2 and τ^2. Figures 1 and 2 show the behavior of the confidence interval for σ^2 as a function of the sample sizes and the relative magnitude of σ^2 to τ^2. The z-axis represents the reduction in the confidence interval relative to the maximum interval width when **u** is not used. Note that substantial reductions in the confidence interval can be made with relatively small n_2, and that the greater the magnitude of τ^2 the more there is to be gained by including the **u** (Figure 1). Also, as the ratio of measurement variance to process variance increases, the value of the auxiliary data sample size, n_2 increases relative to n_1 (Figure 2).

3 Estimating Parameters for a Random Gompertz Time Series in the Presence of Measurement Error

3.1 Model description

Consider a discrete time population growth model, $f(N)$, indexed by time, t, such that:

$$f(N_t) = N_{t+1} = N_t g(N_t) e^{\sigma z_t}. \tag{10}$$

Where N_t is the size of the population at time, t; $g(N_t)$ is a function describing the population growth rate as a function of t; Z_t is a 0 mean random variable and σ^2, process variance, is a scaling term for the variance of Z. If $g(N_t)$ is of the form:

$$g(N_t) = e^{a+b\ln(N_t)}, \tag{11}$$

where b is strictly negative, then the population exhibits density dependent growth with decreasing growth rate as N increases, and negative growth rate above the threshold, $K = e^{-\frac{a}{b}}$ (Gompertz growth model).

Let Y_t be a measurement of the population size at time t. Assume that the measurements are imperfect such that,

$$Y_t = N_t e^{\tau w_t}, \tag{12}$$

where W_t is a random variable and τ^2, measurement error, is a scaling term for the variance of W_t.

If y_t is the natural logarithm of Y_t we can construct the observed population growth rate, r_t, as the difference, $y_{t+1} - y_t$.

$$
\begin{aligned}
\ln(Y_{t+1}) &= \ln(N_{t+1}) + \tau w_{t+1} \\
&= \ln(N_t) + a + b\ln(N_t) + \sigma z_t + \tau w_{t+1} \\
&= \ln(Y_t) - \tau w_t + a + b(\ln(Y_t) - \tau w_t) + \sigma z_t + \tau w_{t+1} \\
y_{t+1} &= y_t + a + by_t - b\tau w_t - \tau w_t + \sigma z_t + \tau w_{t+1} \\
y_{t+1} - y_t &= a + by_t - b\tau w_t - \tau w_t + \sigma z_t + \tau w_{t+1} \\
r_t &= a + by_t - (1+b)\tau w_t + \sigma z_t + \sigma w_{t+1}.
\end{aligned}
$$

If W_t and Z_t are assumed to be independent and identically distributed standard normal random variables, and W_t is independent of W_{t+1}, and Z_t is independent of Z_{t+1}, then r_t, given y_t, is distributed $\mathrm{N}(a + by_t, (1+b)^2\tau^2 + \tau^2 + \sigma^2)$. This statement assumes that the Markov property holds so that the transition r_t is independent of $y_{t-1}, \ldots, y_0$. In practice this assumption is likely to be violated as many biological processes exhibit time lags. For example, the per capita forage resources available when an herbivore population is in decline may be significantly less than that available to an increasing population of the same size due to the time required of plants to grow following intense herbivory. For the sake of simplicity; however, I will assume that the time scale on which the time series is measured is great enough to allow replenishment of resources so that the growth rate is dependent only on the current population size. Conditioning on y_0, the log of the joint likelihood of a, b, σ^2, τ^2, for a time series of length $k+1$ is,

$$
\ln L(a, b, \sigma^2, \tau^2 | \mathbf{r}, \mathbf{y}) = \frac{-k}{2} \ln(2\pi) - \frac{k}{2} \ln[[(1+b)^2 + 1]\tau^2 + \sigma^2] - \frac{1}{2} \frac{\sum_{t=1}^{k} (r_t - a - by_t)^2}{[(1+b)^2 + 1]\tau^2 + \sigma^2}, \quad (13)
$$

where $\mathbf{r}$ is $(r_1, \ldots, r_k)'$ and $\mathbf{y}$ is $(y_1, \ldots, y_k)'$.

3.2 Maximum Likelihood Estimation of σ^2

The parameters σ^2 and τ^2 in equation 13 are confounded and unique maximum likelihood solutions do not exist. Estimates can only be formed for the sum of σ^2 and τ^2. Thus, estimates of the process variance in the time series will always be confounded with the measurement error and yield biased estimates of the process variance.

There are several ways to address the problem of estimating the process variation. First, one could assume the measurements are taken without error. Second, one could provide an independent estimate of τ^2 and pass it into the likelihood as a known quantity (a pseudo-likelihood approach). Lastly, one could form a product of the joint likelihood, $L(a, b, \sigma^2, \tau^2|\mathbf{r}, \mathbf{y})$, and the likelihood for an independent, auxiliary data vector, $\mathbf{u}$, dependent on on τ^2 only, $L(\tau^2|\mathbf{u})$. Both the pseudo-likelihood and full likelihood methods result in unique estimates for σ^2. For simplicity I will consider the case in which the parameters a and b are known and focus on the estimation of the variance components σ^2 and τ^2.

3.2.1 Method 1: "Error Free" Measurement

This approach is appealing for its simplicity, and is common in practice. However, the estimates of σ^2 are inherently biased, with the scale of the bias dependent on the scale of the measurement error.

In this case the measurements are assumed to be taken without error, so that $y_t = \ln N_t$ and $r_t \sim N(a + by_t, \sigma^2)$. In this setting the maximum likelihood estimator of σ^2 is :

$$\hat{\sigma^2} = \frac{1}{k}\sum_{t=1}^{k}(r_t - a - by_t)^2. \tag{14}$$

If; however, there is measurement error, the $\hat{\sigma^2}$ in (14) estimates the quantity $[(1+b)^2+1]\tau^2 + \sigma^2$.

3.2.2 Method 2: Pseudo-Likelihood

If τ^2 is assumed known, the maximum likelihood estimator of σ^2 is:

$$\hat{\sigma^2} = \frac{1}{k}\sum_{t=1}^{k}(r_t - a - by_t)^2 - [(1+b)^2+1]\tau^2. \tag{15}$$

3.2.3 Method 3: Full Likelihood

A logical source of auxiliary data might be a series of log population counts, z_t and $z_{t+\delta}$, made independently of the time series, $\mathbf{y}$, using the same methodology, where δ is an increment short enough to allow the assumption that the population size has not changed. The difference in these log counts $u_t = z_t - z_{t+\delta}$ will be distributed $U \sim N(0, 2\tau^2)$. For a sample of m such repeated log counts, $\mathbf{u} = (u_1, \ldots, u_m)'$, the log of the full likelihood is the product of the likelihood for the time series and the likelihood for the auxiliary data,

$$\begin{aligned}\ln L(a, b, \sigma^2, \tau^2 | \mathbf{r}, \mathbf{y}, \mathbf{u}) &= \frac{-(k+m)}{2} \ln(2\pi) - \frac{k}{2} \ln[[(1+b)^2+1]\tau^2+\sigma^2] - \\ &\frac{1}{2} \frac{\sum(r_t - a - by_t)^2}{[(1+b)^2+1]\tau^2+\sigma^2} - \frac{m}{2}\ln(2\tau^2) - \frac{1}{2}\frac{\sum u_i^2}{2\tau^2} \end{aligned} \quad (16)$$

When a and b are assumed known, the maximum likelihood estimators for τ^2 and σ^2 are,

$$\hat{\tau^2} = \frac{1}{2m} \sum_{i=1}^{m} u_i^2 \quad (17)$$

$$\hat{\sigma^2} = \frac{1}{k} \sum_{t=1}^{k} (r_t - a - by_t)^2 - [(1+b)^2+1]\frac{1}{2m} \sum_{i=1}^{m} u_i^2 \quad (18)$$

3.3 Confidence intervals for σ^2

Confidence intervals can be solved for each of the above cases by inverting the likelihood ratio test statistic,

$$\Lambda(\mathbf{y}) = \frac{L(\sigma_0^2 \mid \mathbf{y})}{L(\hat{\sigma^2} \mid \mathbf{y})} \quad (19)$$

where $\hat{\sigma^2}$ is the maximum likelihood estimate of σ^2. For large sample sizes, $-2\ln\Lambda \sim \chi^2_p$ where p is the number of parameters to be estimated, so an approximate confidence interval can be calculated using the appropriate χ^2 critical value.

3.4 Comparison of Confidence intervals for methods 1-3

I wrote a program in MATLAB to find numerical solutions for the bounds of an approximate 95% confidence interval for σ^2, when a and b are assumed known, using the constrained optimization function 'fmincon' [4]. The routine solved for the maximum and minimum value of σ^2 subject to the constraint,

$$\chi^2_{2,.95} = -2(\ln L(\sigma_0^2 \mid \mathbf{r}, \mathbf{y}, \mathbf{u}) - \ln L(\hat{\sigma^2} \mid \mathbf{r}, \mathbf{y}, \mathbf{u})), \quad (20)$$

which is equivalent to equation 19.

I investigated the effect of the strength of density dependence, magnitude of measurement variance, and sample size of auxiliary data on the performance of the confidence interval and point estimates of the process variance, σ^2 (Tables 1-6). Tables 1-3 present results from time series with weak density dependence ($a = 0.004, b = -0.005$; Figure 3), and Tables 4-6 present results from time series with strong density dependence ($a = 6.4, b = -0.8$; Figure 4). For each density dependence, measurement variance, and sample size combination I generated 1000 time series with measurement error according to equations 10-12 and solved for point and interval estimates of σ^2 using methods 1-3 above. All time series were of length 100 with carrying capacity of approximately 3000 and initial population size of 3000 individuals.

In all scenarios, the maximum likelihood estimates for σ^2 using the "error-free" method were quite biased, and confidence intervals generated using that method rarely contained the true value (Tables 1-6). The 0% coverage probabilities are likely an artifact of the specific parameter values used, but are certainly indicative of poor performance of the estimation method. In all cases the full-likelihood approach resulted in slightly wider confidence intervals, and coverage probabilities that were closer to the nominal 95% rate than the pseudo-likelihood based intervals. This fits the concerns of Davidian and Carroll [5] that pseudo likelihood may underestimate the variance in the estimator.

Surprisingly, despite the difference in the behavior of the time series (Figures 3 and 4), the strength of density dependence does not appear to have a strong affect on the performance of the confidence intervals for σ^2 as the results in tables 1-3 are quite similar to tables 4-6. Coverage rates increased with larger auxiliary data sample sizes, and equaled or exceeded the nominal rate with sample size of 30 when the magnitude of measurement variance was equal to the process variance. Coverage probabilities tended to be slightly lower when the measurement variance exceeded the process variance.

3.5 The estimation of the growth rate parameters of a Gompertz population growth model in the presence of measurement error

Clearly the assumption that the growth rate parameters, a and b, are known is unrealistic. Much debate has focused on statistical tests for density dependence in population time series (i.e. the sign of b). Many methods have been proposed, though their performance is often dependent on the choice of population model and the various tests rarely agree [7]. The estimation of the parameters, a and b, has received comparatively little attention, though is quite important in such applications as predicting extinction risk. A common method for estimating the parameters, a and b, of a Gompertz population growth model has been to regress r_t on $\ln(N_t)$, where $r_t = \ln(N_{t+1}) - \ln(N_t)$, and N_t is the population size at time,t, measured without error [7, 13]. Then $\hat{a}$ is the intercept and $\hat{b}$ the slope of that regression.

In the absence of measurement error ($\tau^2 = 0$) the MLEs of a and b from the likelihood given in equation 13 reduce to the above regression estimates [7]. The MLEs show a negligible bias for short time series and appear to be asymptotically unbiased. However, because $r_t's$ are not independent, confidence intervals based on standard regression results are not valid [7].

In the presence of measurement error ($\tau^2 > 0$), the maximum likelihood estimates of a and b from equation 13 also reduce to the above regression estimates, but the measurement error can introduce significant bias (Figure 5). Interestingly the estimate of the ratio $-\hat{a}/\hat{b}$ is unbiased for $-a/b$, the natural log of the carrying capacity. I present here only results on the bias in $\hat{b}$, noting that they imply a similar, but opposite in sign, bias in $\hat{a}$. The bias in $\hat{b}$ is greatest when the strength of density dependence is weak and is positively related to the magnitude of the measurement error . The bias in $\hat{b}$ is smallest in time series that contain observations far from the carrying capacity (Figure 6).

4 Conclusions and Comments

Measurement error in the monitoring of biological populations is a necessary evil. However, the solution to measurement error need not be restricted to assumptions that "measurement error is negligible", or that the model predictions are "conservative". Neither statement is necessarily true. Population time series measured with error do result in overestimates of environmental variance which will result in higher predicted extinction risk. While this would result in more conservative management and conservation, the overestimation of environmental variation is not the only consequence of measurement error. At least in the Gompertz population growth model, measurement error can result in quite biased maximum likelihood estimates of growth rate and density dependence parameters. Populations with low growth rate and weak density dependence will tend to have positively biased estimates of the growth rate and strength of density dependence. Ginzburg et al. [9] showed that density dependence can, in fact, result in reductions in extinction risk. Thus overestimates of the strength of density dependence may lead to under-estimates of extinction risk, or time to extinction, confounding the argument for conservative estimates.

Biological monitoring very often involves a tradeoff between the number of observations made, and the rigor with which each observation is made. Because many interesting ecological phenomena occur on relatively long time scales, we often put the bulk of our monitoring resources into obtaining time series of maximal length. The above results suggest that such monitoring protocols come at a reasonable cost. One solution may be to allocate some effort to estimating the error in the measurement process. Even a relatively small investment in auxiliary data can result in greatly improved estimates of the process variance.

The source of this auxiliary data is a subject for further investigation. Throughout the above discussion I have assumed that the auxiliary data are a random sample from a normal distribution with mean 0 and variance $2\tau^2$, independent of the time series. This is a reasonable assumption if the auxiliary data are differences of repeated measurements, $u_t = z_t - z_{t+\delta}$, made by the same observer, under similar conditions, using the same protocol. In practice this is difficult to achieve as there is frequently an observer bias, and the second measurement may be influenced by results of the previous count (e.g. the observer now knows where to look for individuals). These issues, however, can be minimized through rigorous field protocol. Double counts made by two separate observers at the same time may also provide reasonable auxiliary data assuming the protocol is followed exactly by both observers. Indeed, this method might be preferred if the time series data are collected by multiple observers.

The use of independent, auxiliary data on the measurement error may also be an inefficient use of resources. An auxiliary data set of size 10 would actually involve 20 counts (10 pairs). It may be possible to incorporate double counts, Y_t and $Y_{t+\delta}$, where Y_t is an observed point in time series. This would reduce by half the number of additional measurements needed, but the auxiliary data would no longer be independent of the time series and that dependence would need to be accounted for in the joint likelihood.

These results would have the most direct application in the development of future monitoring programs in which data on the measurement variance can be collected simultaneously with population counts. These results may also be applied to existing monitoring programs if it can be assumed that the measurement process, in particular the measurement variance, has been constant since the beginning of the monitoring program.

References

[1] Begon, M., J.L. Harper and C.R. Townsend. 90. Ecology: Individuals, Populations, and Communities. Blackwell Scientific Publications. Oxford, England.

[2] Berryman, A.A. 1991. Stabilization or regulation: what it all means! Oecologia 86:140-143.

[3] Boyce, M.S. 1992. Population viability analysis. Annual Review of Ecology and Systematics 23:481-506.

[4] Coleman, T.F., M.A. Branch and A. Grace. 1999. Optimization toolbox: for use with MATLAB. The Math Works, Inc. Natick, MA.

[5] Davidian, M. and R.J. Carroll. 1987. Variance function estimation. Journal of the American Statistical Association 82:1079-1091.

[6] Dennis, B. 1989. Allee effects: Population growth, critical density, and the chance of extinction. Natural Resource Modeling 3:481-538.

[7] Dennis, B., Taper, M.L. 1994. Density dependence in time series observations of natural populations: estimation and testing. Ecological Monographs 64:205-224.

[8] Engen, S., O. Baake and A. Islam. 1998. Demographic and environmental stochasticity-concepts and definitions. Biometrics 54: 840-846.

[9] Ginzburg, L.R., S. Ferguson, H.R. Akcakaya. 1990. Reconstructibility of density dependence and conservative assessment of extinction risks. Conservation Biology 4:63-70.

[10] Gong, G. and F.J. Samaniego. 1989. Pseudo maximum likelihood estimation: theory and applications. Annals of Statistics 9:861-869.

[11] Goodman, D. 1987. The demography of chance Extinction. pp 11-34 *in* Soule, editor, Viable Populations for conservation. Cambridge University Press, Cambridge, England.

[12] Hassel, M.P. 1986. Detecting density dependence. Trends in Ecology and Evolution 1:90-93.

[13] Hooten, M.M. 1995. Distinguishing forms of statistical density dependence and independence in animal time series data using information criteria. Ph.D. Dissertation. Montana State University, Bozeman, USA.

[14] Leigh, E.G. 1981. The average lifetime of a population in a varying environment. Journal of Theoretical Biology. 90: 213-239.

[15] Lindsey, J.K. 1996. Parametric Statistcal Inference. Clarendon Press, Oxford, England.

[16] Royama, T. 1977. Population persistence and density dependence. Ecological Monographs 47:1-35.

[17] Stacey, P.B. and M. Taper. 1992. Environmental variation and the persistence of small populations. Ecological Applications 2:18-29.

[18] Strong,D.R. 1986. Density-vague population change. Trends in Ecology and Evolution 1:39-42.

Table 1: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 0.04$, $b = -.005$, process variance $\sigma^2 = 0.005$, measurement variance $\tau^2 = 0.005$ and sample size for auxiliary data m=10.

	"Error free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0148	0.0054	0.0054
MSE($\hat{\sigma^2}$)	0.1015	0.0159	0.0159
Mean CI length	0.0085	0.0074	0.0095
Mean Upper Bound	0.0198	0.0102	0.0118
Coverage Probability	0.00	0.824	0.89

Table 2: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 0.04$, $b = -0.005$, process variance $\sigma^2 = 0.005$, measurement variance $\tau^2 = 0.005$ and sample size for auxiliary data m=30.

	"Error Free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0150	0.0051	0.0051
MSE($\hat{\sigma^2}$)	0.1048	0.0101	0.01011
Mean CI length	0.0086	0.0080	0.0100
Mean Upper Bound	0.0200	0.0101	0.0117
Coverage Probability	0.00	0.91	0.95

Table 3: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 0.04$, $b = -0.005$, process variance σ^2, measurement variance $\tau^2 = 0.008$ and sample size for auxiliary data m=10.

	"Error Free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0209	0.0065	0.0065
MSE($\hat{\sigma^2}$)	0.2637	0.0344	0.0344
Mean CI length	0.0120	0.0099	0.0126
Mean Upper Bound	0.0280	0.0130	0.0152
Coverage Probability	0.00	0.82	0.93

Table 4: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 6.4$, $b = -.8$, process variance $\sigma^2 = 0.005$, measurement variance $\tau^2 = 0.005$ and sample size for auxiliary data m=10.

	"Error free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0152	0.005	0.005
MSE($\hat{\sigma^2}$)	0.0295	0.0066	0.0066
Mean CI length	0.0059	0.0056	0.0071
Mean Upper Bound	0.0137	0.0085	0.0096
Coverage Probability	0.003	0.87	0.91

Table 5: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 6.4$, $b = -.8$, process variance $\sigma^2 = 0.005$, measurement variance $\tau^2 = 0.005$ and sample size for auxiliary data m=30.

	"Error free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0102	0.005	0.005
MSE($\hat{\sigma^2}$)	0.0293	0.0039	0.0039
Mean CI length	0.0059	0.0058	0.0078
Mean Upper Bound	0.0137	0.0085	0.0096
Coverage Probability	0.001	0.93	0.97

Table 6: Summary of point and interval estimation of σ^2 for 1000 simulations of a Gompertz time series of length 100 with growth parameters $a = 6.4$, $b = -.8$, process variance $\sigma^2 = 0.005$, measurement variance $\tau^2 = 0.008$ and sample size for auxiliary data m=10.

	"Error free"	Pseudo-Likelihood	Full Likelihood
Mean $\hat{\sigma^2}$	0.0134	0.0054	0.0054
MSE($\hat{\sigma^2}$)	0.0738	0.0125	0.0125
Mean CI length	0.0077	0.0069	0.0088
Mean Upper Bound	0.0179	0.0097	0.0114
Coverage Probability	0.000	0.84	0.89

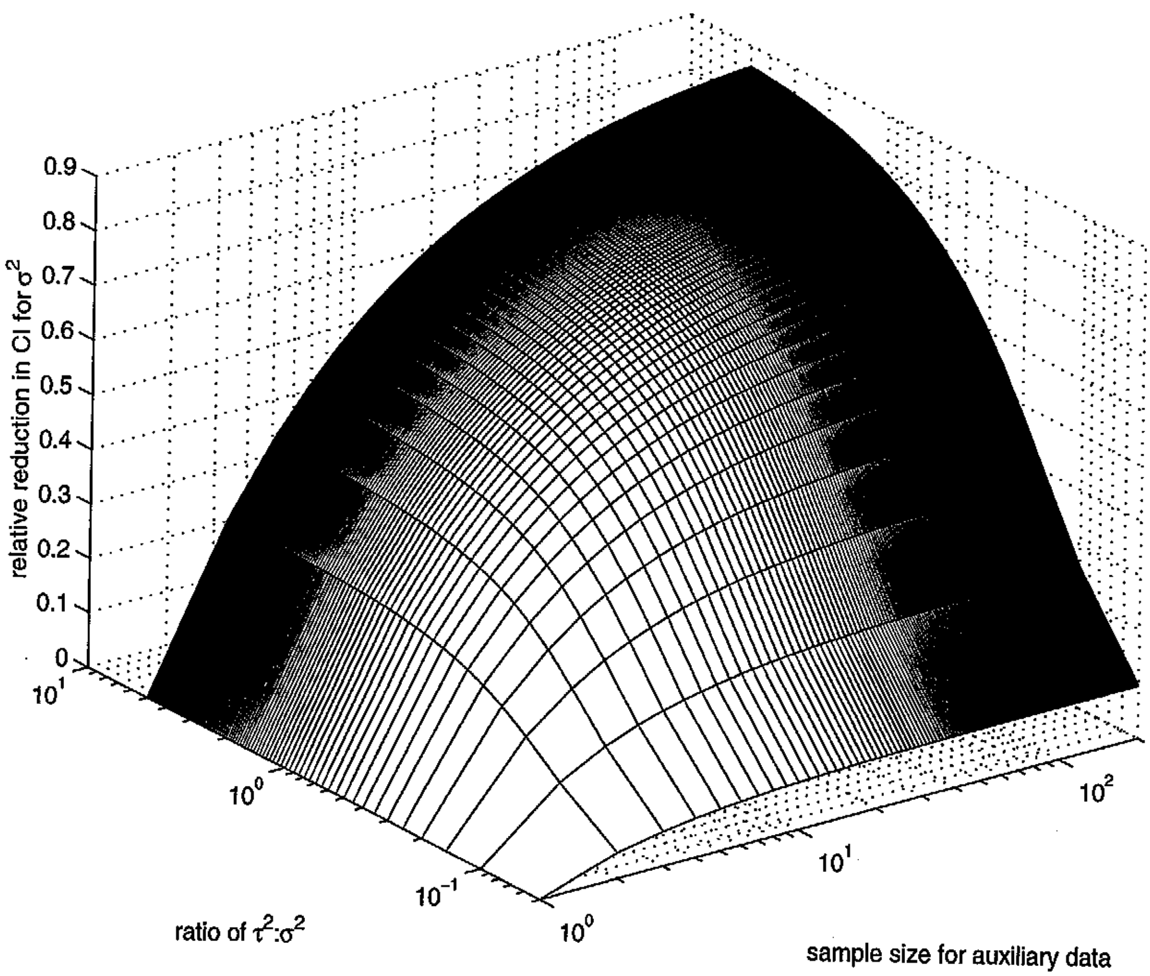

Figure 1: Relative reduction of confidence interval width on σ^2 as a function of the ratio of $\tau^2 : \sigma^2$ and the sample size for the auxiliary information on τ^2. The Z axis is the (max width - width)/ max width.

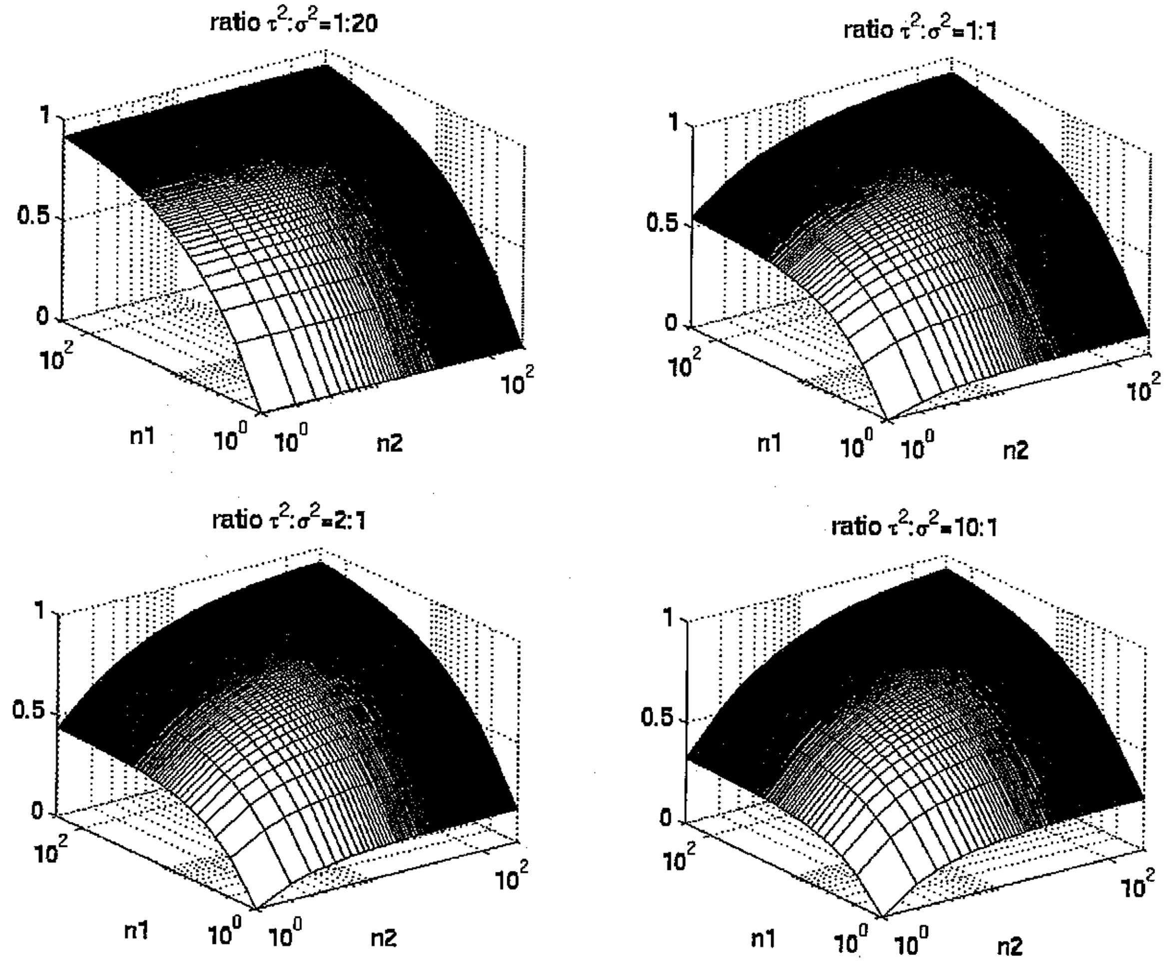

Figure 2: The relative reduction of the confidence interval width on σ^2 as a function of n_1 and n_2 for four ratios of $\tau^2 : \sigma^2$. n_1 is the number of observations of the process, and n_2 is the number of observations of the auxiliary data. The four panels represent increasing magnitude of τ^2 from a ratio of $\tau^2 : \sigma^2$= 1:20 to a ratio of 20:1.

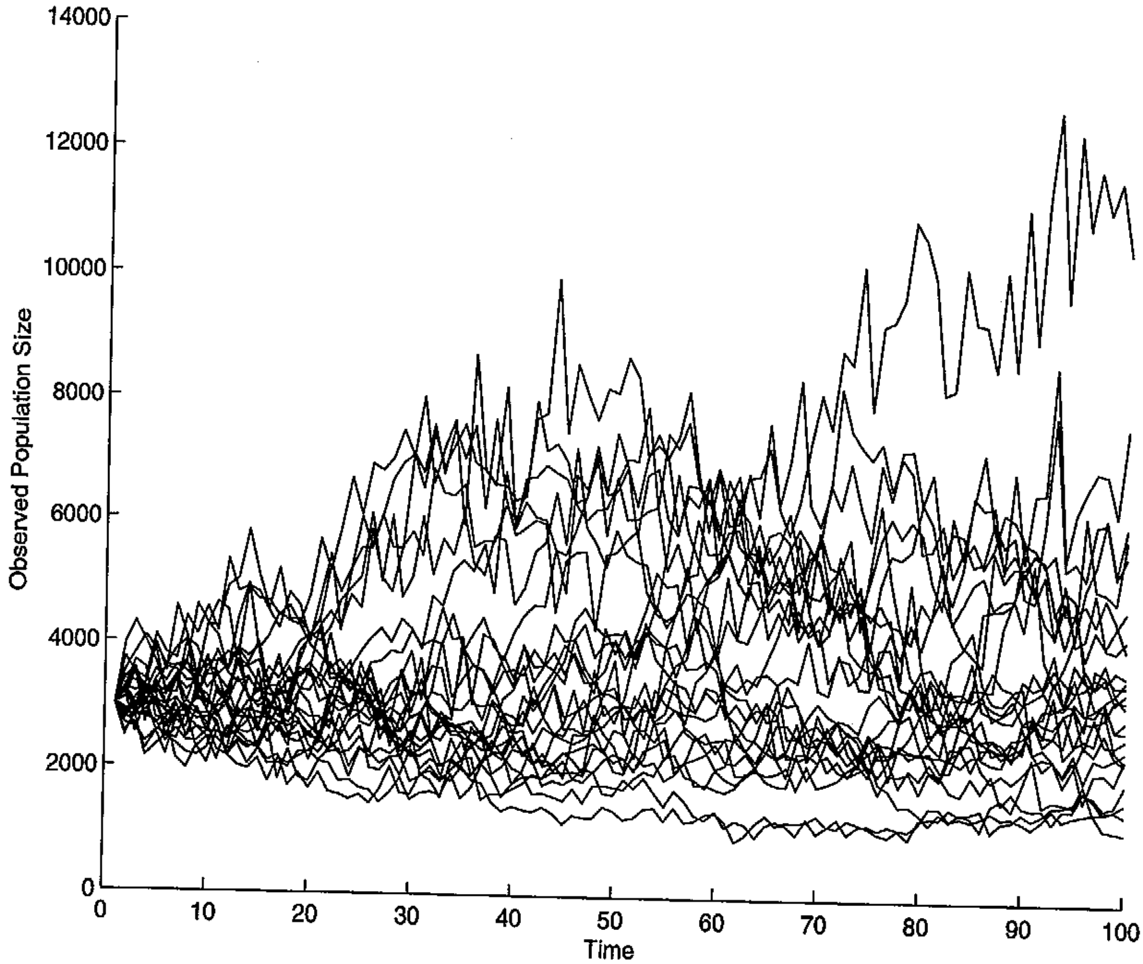

Figure 3: 25 Gompertz time series generated with low population growth rate and weak density dependence ($a = 0.04, b = -0.005$) and $\sigma^2 = 0.005$ and $\tau^2 = 0.005$.

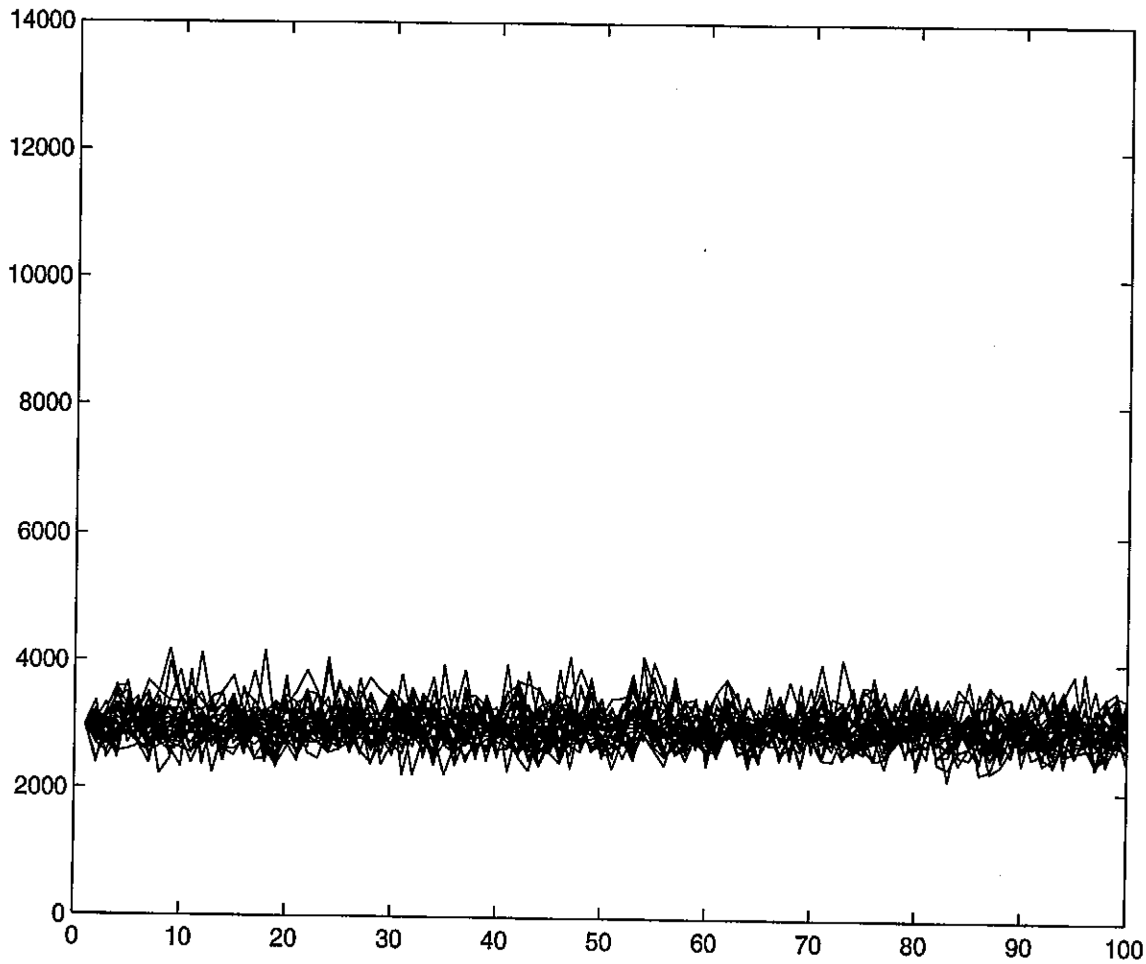

Figure 4: 25 Gompertz time series generated with high population growth rate and strong density dependence ($a = 6.4, b = -0.8$) and $\sigma^2 = 0.005$ and $\tau^2 = 0.005$.

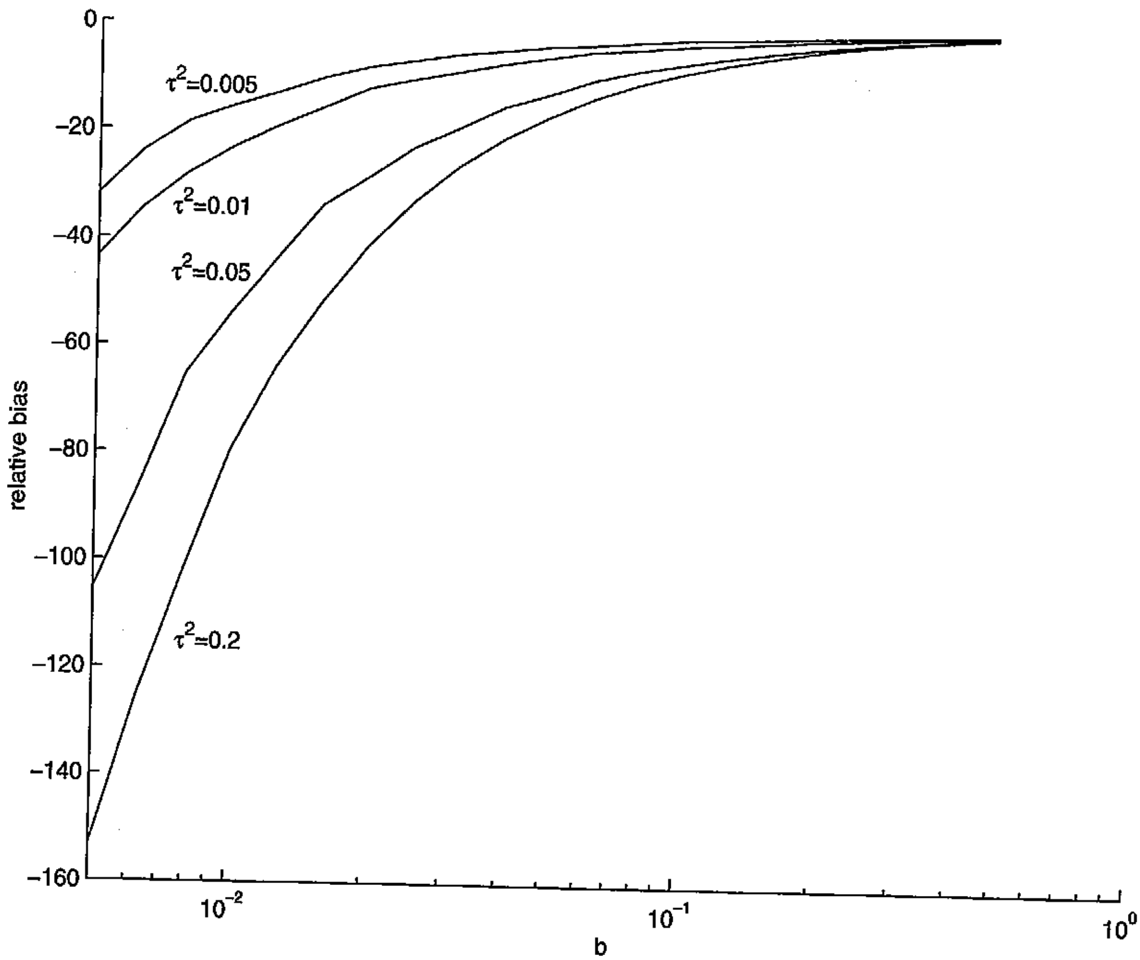

Figure 5: The relative bias in MLE of b as a function b and the measurement variance τ^2, for time series of length 100 with carrying capacity 2981 and $\sigma^2 = 0.005$. Relative bias is $(\hat{b} - b)/b$. The curves for each level represent the mean relative bias for 200 time series at each combination of τ^2 and b.

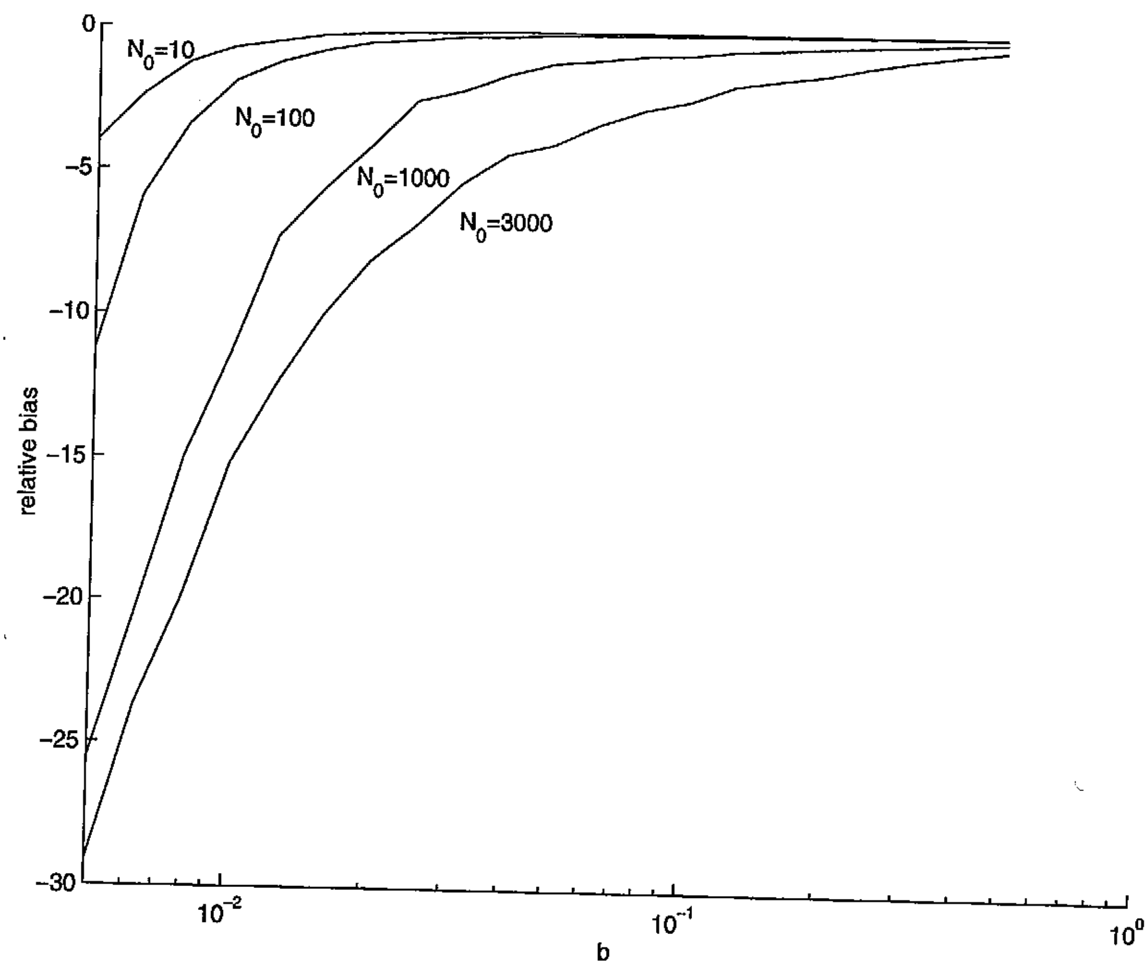

Figure 6: The relative bias in MLE of b as a function b and the initial population size N_0, for time series of length 100 with carrying capacity 2981 and $\sigma^2 = \tau^2 = 0.005$. Relative bias is $(\hat{b} - b)/b$. The curves for each level represent the mean relative bias for 200 time series at each combination of N_0 and b.