Contributed Papers

The Use of Species Accumulation Functions for the Prediction of Species Richness

JORGE SOBERÓN M.
Centro de Ecología
Universidad Nacional Autónoma de México
Apdo. Postal 70-275
México D. F.

JORGE LLORENTE B.
Facultad de Ciencias
Universidad Nacional Autónoma de México
Apdo. Postal 70-399
México D. F.

Abstract: *We develop a stochastic theory of the accumulation of new species in faunistic or floristic inventories. Differential equations for the expected list size and its variance as a function of the time spent collecting are presented and solved for particular cases. These particular cases correspond to different models of how the probability of adding a new species changes with time, the size of the list, the complexity of the area sampled, and other parameters. Examples using field data from butterflies and mammals are discussed, and it is argued that the equations relating sampling effort with size of the list may be useful for conservation purposes because they should lend formality to comparisons among lists and because they may have predictive power by extrapolating the asymptotic size of the lists. The suitability of different models to a variety of field situations is also discussed.*

El uso de funciones de acumulación de especies para la prediccion de la rigueza de especies

Resumen: *Se presenta una teoría estocástica de la acumulación de especies nuevas en inventarios florísticos o faunísticos. Se obtienen y resuelven casos particulares de las ecuaciones diferenciales que relacionan el tamaño esperado de las listas y su variancia con el tiempo dedicado a la colecta. Los casos particulares corresponden a diferentes modelos de cómo la probabilidad de añadir una especie nueva a la lista cambia con el tiempo, el tamaño de la lista, la complejidad del área y otros parámetros. Se discuten ejemplos con datos de campo de mariposas y mamíferos y se argumenta que el contar con ecuaciones que relacionen el esfuerzo de colecta con el tamaño del inventario puede ser útil para propósitos conservacionistas porque se podrán formalizar las comparaciones entre inventarios y porque tales ecuaciones pueden tener un valor predictivo al extrapolar para obtener los valores asintóticos de las listas. También se discute la conveniencia de los diferentes modelos a distintas situaciones de campo.*

Introduction

Faunistic and floristic studies often reveal that as the time spent collecting increases, the number of new spe-

cies added to the list asymptotically approaches some ceiling. In a paper on inventories of butterfly species, Clench (1979) proposed the use of the Michaelis-Menten equation to describe empirically the behavior of the cummulative species-effort relationship. Despite the potential utility of such a relationship, lepidopterists have only recently begun to use it (Lamas et al. 1991;

Raguso & Llorente 1993). Neither Clench's nor other related equations are commonly used in faunistic studies. In at least one botanic paper, Miller and Wiegert (1989) have used a related equation (an exponential model; see below) to predict the total number of plant species expected in a region. Although the use of such functions is still uncommon, it is more widespread in plotting species versus effort to estimate visually whether an asymptote has been reached (Miller & White 1986; Miller et al. 1987; Miller & Wiegert 1989; Newmark 1991).

Having a theoretical basis for understanding the relationship between collecting time and number of species accumulated would be useful because, among other things, (1) it would give formality to faunistic and floristic work by allowing more rigorous and quantitative comparisons between lists, (2) it would provide a planning tool for collecting expeditions, and (3) it may provide a predictive tool for conservation and biodiversity studies, if used to extrapolate the total number of species present in an area. In this paper, we present a stochastic model of the process of adding new species to a list and we will derive solutions for different biological situations. For one of these we obtain the variance, the lacking of which, as pointed out by Lamas et al. (1991), is a drawback of Clench's model. We fit a number of data sets to our equations and discuss the usefulness and limitations of this method.

The Model

A simple model of the process of accumulating new species is the pure birth process (Bailey 1964; Pielou 1969). This model assumes that the system is represented by states, in our case the number of different species, and that a suitable time increment may be chosen such that the system either moves to the next state or remains where it was at time t. In symbols:

$$\text{prob}(j \to j + 1)_{\Delta t} = \lambda(j,t) \, \Delta t \qquad (1)$$
$$\text{prob}(j \to j)_{\Delta t} = 1 - \lambda(j,t)\Delta t.$$

In words, the probability of adding one species to a list of size j in the time interval Δt is denoted by $\lambda(j,t)\Delta t$, and we assume the time interval is so small that the only other possibility is that in the same time interval no new species is found, with probability $1 - \lambda(j,t)\Delta t$. The symbol $\lambda(j,t)\Delta t$ denotes the probability of adding a new species to the list, after a collecting time Δt and given that we already have j species in time t. The expression $\lambda(j,t)\Delta t$ is a per-unit time *transition* probability. Henceforth we shall refer to $\lambda(j,t)$ as the collecting function. It should be clear that the particular shape of $\lambda(j,t)$ depends on factors such as the sampling method, the size of the area sampled, and coverage of suitable habitats.

The collecting function is a function both of the biology of the taxon of interest and of the methods used.

With this definition, we now ask for the probability $p(j)_t$ that at time t the list has exactly j species (this is a *state* probability). It is shown in text books of stochastic processes that such a probability obeys the following equations:

$$d\,p(j)_t/d\,t = p(j - 1)_t \, \lambda(j - 1,t) - p(j)_t \, \lambda(j,t). \qquad (2)$$

It is important to make the technical point that since the only permitted transition is from $j - 1$ to j in the time unit Δt, Equations 2 are a particular case of the nonhomogeneous and more general Kolmogorov's forward equations (Bailey 1964: 77), and it is allowed to have λ as a function of time.

The above set of equations (one for each state j), if solved, will yield the distribution of probabilities of the size of the list at time t. Although the system can be solved for particular models $\lambda(j,t)$, and purpose of this paper requires only expressions for the expected size and variance of the list. After some algebra (outlined in the appendix) it can be shown that the differential equations for the first and second moments of the distribution of j at time t are simply

$$d\,\langle j\rangle/d\,t = \Sigma \, p_t j \, \lambda(j,t) \qquad (3)$$

$$d\,\langle j^2\rangle/d\,t = 2\,\Sigma\, j\, p_t(j)\, \lambda(j,t) + \Sigma\, p_t(j)\, \lambda(j,t), \qquad (4)$$

where the sums are taken from $j = 0$ to infinity. Generally speaking, after substitution of particular models of the collecting function $\lambda(j,t)$ in Equations 3 and 4, we solve the differential equations and obtain the expected value of the number of species in time t: $E(j,t) = \langle j\rangle_t$, henceforth denoted as $S(t)$, and its variance $V(j,t) = \langle j^2\rangle_t - \langle j\rangle^2_t$. A number of interesting quantities can, in principle, be derived from the moments. For example, the list size for long times, the time to accumulate a certain fraction of the asymptote, the time to lower the per capita rate of species increase below a certain threshold, and the confidence limits follow from the solutions to Equations 3 and 4. In the following section we shall find some of these for particular cases.

Linear Dependence on j

The simplest case is when the collecting function depends linearly on the size of the list and the parameters are constant in time:

$$\lambda(j,t) = a - b\,j, \qquad (5)$$

meaning that as the species list grows, the probability of adding a new species to the list in the interval Δt decreases proportionally to the current size of the list, eventually reaching zero. This model may be adequate

when one is sampling a relatively small area, or a well-known group, or both, and eventually all species will be registered.

The expressions for the mean and variance obtained by substituting Equation 5 in Equations 3 and 4 and solving the differential equations are

$$S(t) = a/b[1 - \exp(-b\,t)] \qquad (6)$$

$$V(t) = S(t)\exp(-b\,t). \qquad (7)$$

The parameter a represents the list increase rate at the beginning of the collection, and the asymptote is given by a/b. a has units of species $\times$ time^{-1}, and b of time^{-1}. Both parameters can be obtained by nonlinear regression procedures, and below we shall assume that a suitable algorithm is available without entering into the details of nonlinear fitting.

Lamas et al. (1991) asked for the time t_q required to register a proportion of the total fauna $q = S/R$, where $R = a/b$ represents the asymptote or total richness of the site. From Equation 6, this is simply

$$t_q = -1/b\ln(1 - q). \qquad (9)$$

For example, how long it will take to reach 90% ($q = 0.9$) of the asymptotic size of a list if $b = 0.1$ per week? Equation 9 gives the answer as 23 weeks. Although this result is interesting, given the asymptotic behavior of Equation 6, reaching a 100% richness requires an infinite time. It may be more useful to ask how long it will take for the rate of per capita species increment (dS/Sdt) to go below a particular size. For example, how much collecting time it is required for $dS/Sdt < 0.01$? Calling the threshold k (which has units of time^{-1}), the time needed to lower the per capita list increase rate is simply:

$$t_k = 1/b\ln(1 + b/k). \qquad (10)$$

If, for example, $b = 0.1$ per week, and $k = 1\%$, then $t_k \simeq 24$ weeks, meaning that 24 weeks after the beginning the list will be growing at 1% of the current size, per week.

Since the standard error of $S(t)$ is the square root of the variance, we can use Equation 7 to estimate confidence limits of a given species count. In particular, we can estimate confidence limits for $S(t_k)$, the number of

species collected at time t_k. Substituting Equation 10 in Equation 6, we obtain $S(t_k) = a/(b + k)$, and the corresponding standard error is $(a\,k)^{1/2}/(b + k)$. In the absence of a full probability distribution for $S(t)$, it is possible to use the rule of thumb that two standard errors approximate a 95% confidence interval. The above results are summarized in Table 1.

Exponential Dependence on j.

A slightly more complex model for the collecting function arises when we assume that increasing the size of the collection decreases the probability of adding a new species in a nonlinear way. The simplest supposition is an exponential decrease:

$$\lambda(j,t) = a\exp(-b\,j). \qquad (11)$$

This model may be reasonable in cases in which the region being sampled is large or the taxa poorly known, and thus the probability of finding a new species never reaches zero.

Substitution of Equation 11 in Equation 3 yields an equation that can be solved by noting that $\Sigma\,p(j)$ $\exp(-b\,j)$ is the definition of the probability-generating function (PGF) of the distribution $p(j)$. By postulating different distributions, we can solve the equation. A reasonable assumption is that the $p(j)$ are Poisson distributed, with $PGF = \exp(-z\langle j\rangle)$ and $z = 1 - \exp(-b)$. Then it is possible to obtain the expectation $S(t)$, which is simply:

$$S(t) = 1/z\ln(1 + zat).$$

Another complication arises when we have exponentially decreasing probabilities of adding a new species, but allow them to reach a value of zero:

$$\lambda(j,t) = a\exp(-b\,j) - c. \qquad (12)$$

Again, the expectation $S(t)$ can be obtained:

$$S(t) = 1/z\ln[a/c - (a - c)\exp(-czt)/c].$$

The Clench Equation

The Michaelis-Menten equation used by Clench (1979) can be derived from the model presented here, by going

Table 1.   Some statistics of the collecting functions discussed.

	Exponential	Logarithmic	Clench
$\lambda(j,t)$	$a - bj$	ae^{-bj}	$a + b/a[S(t)^2 - 2j$
$S(t)$	$ab(1 - e^{-bt})$	$1/z\ln(1 + zat)$	$at/(1 + bt)$
$Var(t)$	$S(t)e^{-bt}$		
t_q	$1/b\ln[1/(1 - q)]$		$q/[b(1 - q)]$
t_k	$1/b\ln(1 + b/k)$	$k \geq za/[(1 + zat)\ln(1 + zat)]$	$[(1 + 4bk)^{1/2} - 1]/2b$
Asymptote	a/b		a/b

backwards on the derivation to obtain its implicit collecting function:

$$\lambda(j,t) = a + b^2/a \, S(t)^2 - 2 \, b \, j/a \qquad (13)$$

or, equivalently,

$$\lambda(j,t) = a + b^2/a \, [a \, t/(1 + b \, t)]^2 - 2 \, b \, j/a \qquad (14)$$

Substituting either of the above equations in Equation 3 and solving yields

$$S(t) = a \, t/(1 + b \, t),$$

which is Clench's equation with a slightly different parametrization. In Equation 13, the expectation appears in the collecting function, thereby increasing its value. Similarly, in Equation 14, for a given value of j the collecting function is larger if the time accumulated is larger. Biologically, this means that the probability of adding new species will improve (up to a ceiling) as more time is spent in the field. This seems to be a very plausible mechanism. It makes sense to suppose that as one accumulates experience with the site, taxa, and methods, the chances of adding new species will improve. It is very interesting that Clench's equation, originally proposed only on empirical grounds, appears to have a sensible theoretical basis.

Other particular cases can be solved: for example, an exponential collecting function with negative-binomial distribution of the $p(j)$s, and some time-varying functions. Clearly, each set of assumptions about collecting functions will yield different predictions of the size of the species accumulation in inventory studies. We will proceed to fit some data sets and to discuss the results.

Examples

Lamas et al. (1991) present data obtained from a 200-person-hours collection (during September 1989) in the Pakitza biological station, Parque Nacional Manu, Madre de Dios, Peru. They fit their data to the equation of Clench and obtained a very good fit. They also estimated the asymptote (905 species) and calculated the time required to reach different percentages of it. We digitized the information from their Graph 1, and in Figure 1 and Table 2 we present the results of fitting the Clench, the exponential, and the logarithmic functions to data from Lamas et al. (1991). The models were fitted by the quasi-Newton method provided by the package STATISTICA (StatSoft 1991).

It is clear that although the data fit well to each of the functions, they extrapolate to very different numbers of species. In fact, it is impossible to choose the best of the three models based solely on the data set. To choose an equation, one has to decide which underlying collecting model describes most accurately the particular situation. For the three equations discussed above the models are (1) the probability of adding new species decreases linearly with the size of the list (Equation 5); (2) adding a new species becomes more and more difficult, but never reaches zero (Equation 11); and (3) the probability of adding a new species eventually vanishes, but field experience increases it (Equation 13).
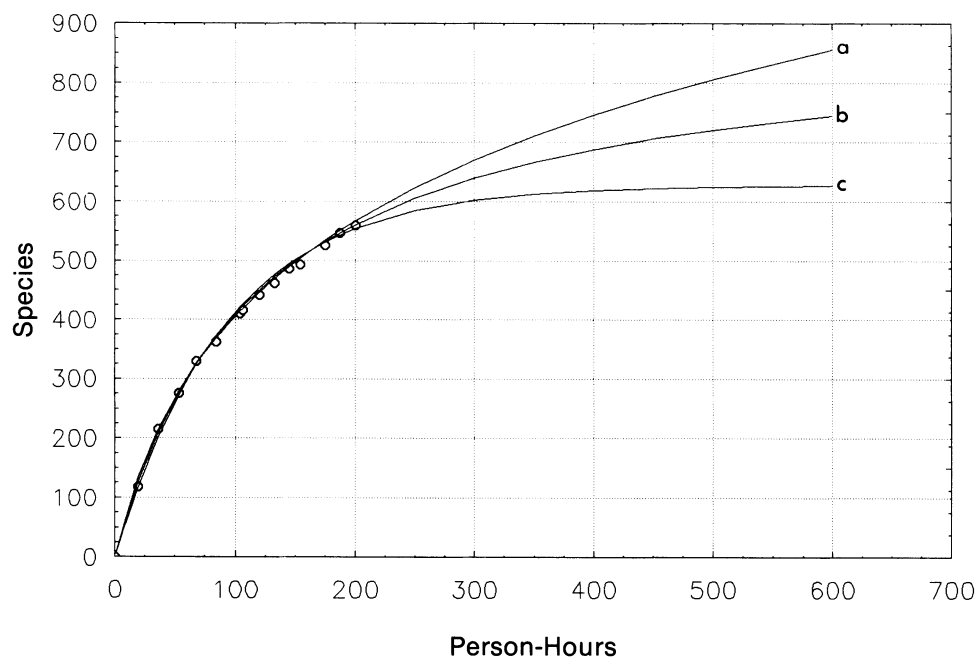

Figure 1. Accumulation curve for butterflies in Pakitza, Peru. Data from Lamas et al. (1991). a corresponds to the logarithmic equation, b to Clench's equation, and c to the exponential equation.

Table 2.   Regression statistics of the four examples.

		Pakitza[1]	Atoyac[2]	Chajul[3]	Powdermill[4]
	r^2	0.99	0.96	0.967	0.986
Clench	a	7.57	6.64	2.88	1.073
	b	0.0085	0.0134	0.035	0.0135
	asymptote	890	495	82	79
	r^2	0.99	0.95	0.972	0.988
Exponential	a	6.72	5.63	2.65	0.761
	b	0.011	0.0155	0.047	0.011
	asymptote	611	363	56	69
	r^2	0.99	0.97	0.96	0.917
Logarithmic	a	8.869	8.413	3.207	2.447
	z	0.00347	0.00675	0.0356	0.065

Time units:
[1] *person-hours*$^{-1}$.
[2] *person-days*$^{-1}$.
[3] *nights*$^{-1}$.
[4] *person-hours*$^{-1}$.

In the case of the data of Lamas et al. (1991), it seems likely that the log model (Equation 11) will be adequate to extrapolate, given the size of the area, the complexity of the fauna, the fact that the list size is still far from the asymptote, and the yearly fluctuations many tropical butterfly species undergo. All these points suggest that the probability of finding new species will still be different from zero after a sizeable increase of the collection effort. It is interesting to extrapolate the models fitted to the first 200 hours of data to the list size that Robbins (personal communication) has reported after 565 person-hours. At that time the Clench model predicts 737 species and the log model 839, while the true

value is 979. Although this extrapolation covers an increase of more than 180% over the time interval used for fitting the models, and this interval includes a non-asymptotic part of the curve, the log model predicts the correct value within 15%.

In another case, Vargas et al. (1991) reported their butterfly sampling, over three years, of a large transect (300–2500 meters above sea level) from semideciduous rain forest to pine forest. In Table 2 and Figure 2, we present the results of fitting the models. As in the Pakitza data, the three models provide excellent fittings in terms of explained variance, display similar residual distributions, but predict contrasting long-term behavior.
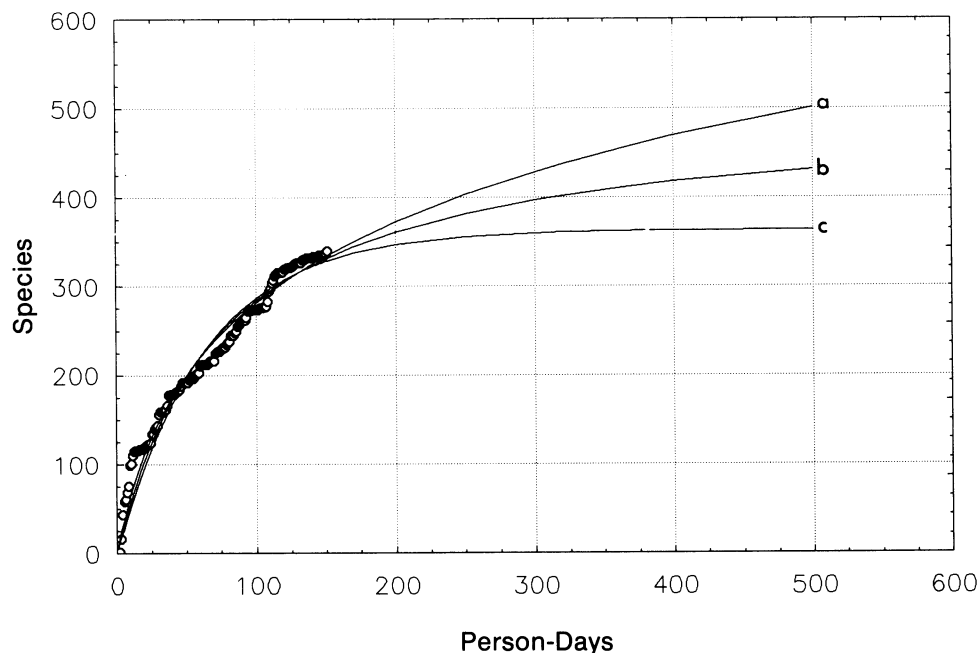

Figure 2. Accumulation curve for butterflies in Sierra de Atoyac, México. From Vargas et al. (1991). a corresponds to the logarithmic equation, b to Clench's equation, and c to the exponential equation.

As in the previous example, and for similar reasons, it is reasonable to assume that either the log or Clench's model may be better predictors of the future behavior of the sampling effort. They predict a species increase of around 15% (Clench) or 20% (log) if sample effort is doubled to 300 person-days. The exponential model, on the other hand, predicts an increase of only about 1% after doubling the effort, which in this case seems unlikely small.

Another example is the list of bat species reported by Medellín (1986 and unpublished) from the Chajul Biological Station in the Lacandon rainforest of southern Mexico. The collections of bats have been ongoing for about seven years, using mist nets at several spots near the station. Table 2 and Figure 3 show the results. As in the previous cases, variance explained by each model and residuals are very similar. According to Medellín (personal communication), his methods are well established and the area, although very rich, is relatively small (a few hundred hectares), so the exponential model should apply. This predicts an increase of about 10% after doubling the sampling time.

In our last example we reanalized the example provided by Clench (1979) to illustrate his species-time formula. This study was carried out for 13 years and totaled 820 hours of collecting and observing butterflies at the 2000-acre Powdermill Nature Reserve in Westmoreland County, Pennsylvania. The list appears to be almost in the asymptote, with only one species added in the last four years of data. The results appear in Table 2 and Figure 4. The nonlinear fit to Clench's model gives

an asymptote of 79 species. Clench (1979) did not specify his fitting method, which yields an asymptote of 78. However, he suggests a simplified method based on eye-fitting a curve to the data. This is unreliable, as we have seen that very different predictions can be obtained from fitting a variety of models. By doubling the sampling time and using Clench's model, an increase of 4% of the list is predicted. The exponential model, which assumes a linear decrease of the probability of adding a new species, should not be as good a model for the sampling of butterfly fauna because temperate lepidopteran species are known to undergo marked abundance cycles (see Taylor & Taylor 1977), and therefore the probability of adding new species should decrease slower than linearly. The fit of the exponential model to Clench's data illustrates a problem raised by Lamas et al. (1991): the estimated species richness is smaller than the last data point. This is due to the very quick approach to the asymptote that characterizes the exponential model. As Lamas et al. (1991) state, this problem can be overcome by fitting the data with a high weight assigned to the last point.

Discussion

The use of extrapolations of spatial data to estimate species richness is not new (Kernshaw 1973; Palmer 1990, among others) and can be traced back to the classical works of Preston (1948, 1962a, 1962b), Fisher et al. (1948), and others. To our knowledge, however, extra-
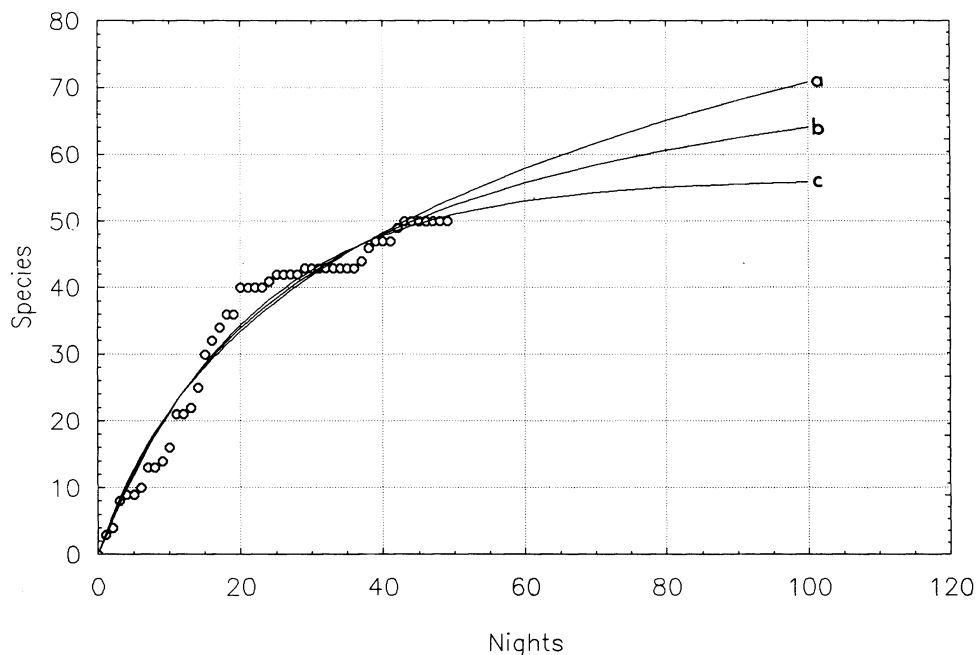

Figure 3. Accumulation curve for bats in Chajul, México. From Medellín (1986 and unpublished). a corresponds to the logarithmic equation, b to Clench's equation, and c to the exponential equation.
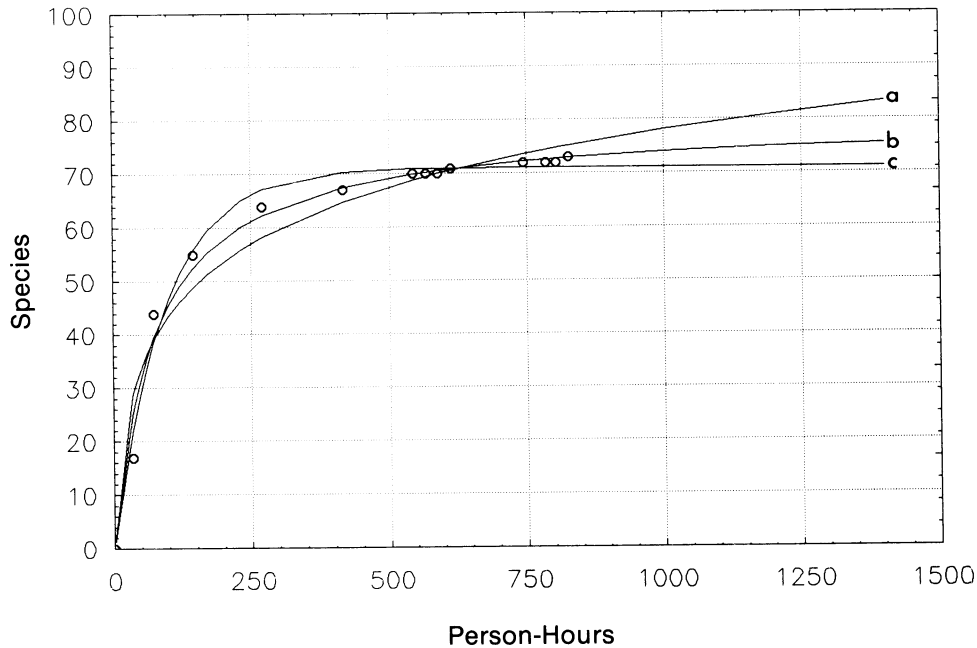
Figure 4. Accumulation curve for butterflies in Powdermill Reserve, Pennsylvania. Data from Clench (1979). a corresponds to the logarithmic equation, b to Clench's equation, and c to the exponential equation.

apolations over the time domain have not been systematically pursued by conservation biologists (Clench 1979; Miller & White 1986; Lamas et al. 1991).

One of the potential uses of such methodology could be to lend rigor to faunal inventories of areas. In poorly collected sites, which often are important for conservation purposes, reporting a number of species may be misleading without some information about how far from complete such lists are. Either the rates of accumulation of new species or an estimate of the percentage of the total number is necessary to make meaningful comparisons. Obviously, a place in which 80 species of butterflies have been reported with 0.1 additional species/person-hour is very different from a place with the same 80 species and a rate of 0.01 additional species/person-hour. In order to make such comparisons possible, the effort (time/person and number of persons) allocated to the addition of new species should be reported. It is clear, however, that as with the size and composition of the list, the effort of persons of different expertise may not be equivalent, thus hindering comparisons between lists. It is also clear that time in itself is not what counts, but how this time is distributed over the seasons. For example 50 person-hours during the dry season may be very different from the same 50 person-hours well distributed over one year. We shall return to this point later.

Predicting the richness of the fauna of a site, given the known accumulation curve, would be interesting. We believe that the models presented here can be used to this purpose with some precautions. First, a sample biased either temporally or spatially is useless for extrap-

olation. For example, collecting only during the rainy season, or only in the understory, edges, or canopy of a forest, will yield extrapolations valid only for the spatial and temporal conditions sampled. The curves aggregate variation in the taxa, the sampling methods, and the spatial and temporal heterogeneity affecting the organisms, and extrapolations must take this into account.

Second, choosing an adequate model of the collecting methods is critical to accurate estimation of faunal size. Different models diverge significantly in their extrapolations while fitting exceedingly well to the same set of data. In this paper we have used three models, but the general theory presented allows the derivation of accumulation curves for a variety of collecting functions. In choosing a suitable model, the researcher needs to state explicitly its underlying assumptions. Because this choice is to some extent subjective, developing more objective procedures for choosing a model should be a priority.

Choosing among the different models requires information about the size of the area sampled and the kind of fauna or flora in question. One extreme case is sampling well-known taxa in small or homogeneous areas with few rare species. In this case, the exponential model may be suitable. The other extreme is sampling unknown taxa in large or heterogeneous areas with many rare species. The Clench or logarithmic models may be adequate for these situations.

Clearly, there must be a relationship between the sampled area, the species-abundance curve, and the collecting function. Several authors have advanced in this direction. For example, Miller and Wiegert (1989) gen-

erated species-abundance relations with canonical log-normal, uniform, random, and observed extant species. Then they obtained the accumulation exponential curves by computer-sampling from these data sets. Both the asymptote (a/b) and the increase rate of list size near the origin (a) appear to be similar for the different species-abundance distributions, but there are differences in the middle part of the accumulation function sampled from different distributions (Miller & Wiegert 1989). From a different point of view, Efron and Thisted (1976) developed a method for the estimation of the number of new species that will appear after sampling a time unit. Unfortunately, their method requires the species-abundance distribution as obtained by sampling the "fauna" during a previous time unit, which is difficult. Finally, an anonymous referee points out that a log-series distribution (Pielou 1969) of species-abundance in which the number of individuals sampled increases linearly with time will yield the logarithmic model (Table 1). This interesting subject presents some difficulties and will be addressed in a future paper.

Another application of the accumulation functions may be the planning of field campaigns. By estimating the number of hours required to add a given number or percentage of species, given a previous history, it should be feasible to estimate costs of field work in a rigorous way. Not only might this make possible the estimation of the cost of adding new species to the list, but because near the asymptote rare species are likely to be the ones being added, it may be possible to obtain some value/cost estimate for different periods during the collection.

All the curves fitted present a very regular distribution of residuals. This indicates systematic departure from the assumptions of regressions. Also, the data points are not independent. These two points, strictly speaking, invalidate statistical inference, but this is a point of statistical finesse that may be irrelevant for the purposes of this paper. Normally, the biologist tends to assess the total richness of a site by extrapolating from his or her experience of the place, methods, and taxa, without assigning any probability of error to the figure. The method presented here is a way to add objectivity and rigor to such informal practices. If only because they expose their hidden assumptions, the methods presented are interesting. More experience with the methodology and further development of the theory—in particular its statistical aspects—will be required to decide whether they are useful for prediction or planning.

Acknowledgments

Literature Cited

Bailey, N. T. J. 1964. The elements of stochastic processes. Wiley Publications in Statistics, New York, New York.

Clench, H. 1979. How to make regional lists of butterflies: Some thoughts. Journal of the Lepidopterists' Society 33(4):216–231.

Efron, B., and R. Thisted. 1976. Estimation of the number of unseen species. Biometrika 63:435–447.

Fisher, R. A., A. S. Corbet, and C. B. Williams. 1948. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12:42–58.

Kernshaw, K. A. 1973. Quantitative and dynamic plant ecology Arnold, London, England.

Lamas, G., R. K. Robbins, and D. J. Harvey. 1991. A preliminary survey of the butterfly fauna of Pakitza, Parque Nacional del Manu, Peru, with an estimate of its species richness. Publicaciones del Museo de Historia Natural, Universidad de San Marcos, Peru 40:1–19.

Medellín, R. 1986. Murciélagos de Chajul. B.Sc. Thesis. Facultad de Ciencias, Universidad Nacional Autónoma de México, México.

Miller, R. I., and P. S. White. 1986. Considerations for preserve design based on the distributions of rare plants in Great Smoky Mountains National Park. Journal of Environmental Management 10:119–124.

Miller, R. I., and R. G. Wiegert. 1989. Documenting completeness, species-area relations, and the species-abundance distribution of a regional flora. Ecology 70(1):16–22.

Miller, R. I., S. Bratton, and P. S. White. 1987. A regional strategy for reserve design and placement based on an analysis of rare and endangered species distribution patterns. Biological Conservation 39:255–268.

Newmark, W. D. 1991. Tropical forest fragmentation and the local extinction of understory birds in the eastern Usamabara Mountains, Tanzania. Conservation Biology 5(1):67–78.

Palmer, M. W. 1990. The estimation of species richness by extrapolation. Ecology 71(3):1195–1198.

Pielou, E. C. 1969. An introduction to mathematical ecology. Wiley-Interscience, New York, New York.

Preston, F. W. 1948. The commonness, and rarity, of species. Ecology 29:254–283.

Preston, F. W. 1962a. The canonical distribution of commonness and rarity, vol. I. Ecology 43:185–215.

Preston, F. W. 1962b. The Canonical Distribution of Commonness and Rarity, vol. II. Ecology 43:410–432.

Raguso, R., and J. Llorente. 1993. The butterflies of the Tuxtlas Mts. Veracruz, México, revisited: Species-richness and habitat disturbance. Journal of Research on the Lepidoptera. In press.

StatSoft. 1991. CSS: Statistica handbook, vol. II. StatSoft Inc., Tulsa, Oklahoma.

Taylor, L. R., and R. A. J. Taylor. 1977. Aggregation, migration and population mechanics. Nature 265:415–420.

Vargas F. I., J. Llorente, and A. Luis. 1991. Lepidopterofauna de Guerrero. I. Distribución y Fenología de los Papilionoidea de la Sierra de Atoyac. Publicaciones Especiales del Museo de Zoología #2. Facultad de Ciencias, Universidad Nacional Autónoma de México, México.

Appendix

To derive Equations 3 and 4, we begin with the definitions of the first two moments:

$$\langle j \rangle = \Sigma\, j\, p(j)_t$$
$$\langle j^2 \rangle = \Sigma\, j^2\, p(j)_t,$$

which have derivatives:

$$d\langle j \rangle/dt = \Sigma\, j\, dp(j)_t/dt \qquad (A1)$$

$$d\langle j^2 \rangle/dt = \Sigma\, j^2\, dp(j)_t/dt. \qquad (A2)$$

Substitution of the values of $dp(j)_t/dt$ given by Equation 2 yields equations that can be simplified, in the case of A1, by adding and subtracting $\Sigma p(j - 1)\lambda(j - 1)$ and then simplifying and, in the case of A2, by adding and subtracting terms to complete the expressions $\Sigma p(j - 1)(j - 1)^2$ and $\Sigma p(j - 1)(j - 1)^3$. Further simplification yields Equation 4.