

Regression Approximations for Continuous-Time, Stochastic Compartmental Models of Discrete Interacting Populations

Richard L. Patterson

*School of Natural Resources
The University of Michigan
Ann Arbor, Michigan 48109*

Transmitted by Melvin R. Scott

ABSTRACT

Canadian lynx and snowshoe hare interact in a predator-prey relationship in which lynx abundance is more strongly dependent upon hare abundance than vice versa. A four-compartment stochastic compartmental model is specified that represents net balances of live individuals in each population, together with cumulative, nondecreasing balances of dead individuals of both populations. Reproduction is modeled as time-variant Poisson arrival processes. Residence times in live compartments, equivalent to ages of individuals, are modeled as conditional distributions, dependent upon abundances of the other population in each case. Models of residence-time distributions are hypothesized for each population and fitted to available data on abundance in live compartments over an 85-year period. Multiple regression techniques are used to fit the models, from which conditional distributions of population abundances in live compartments are used to compute numerical estimates of net balances from year 1850 to year 1935 and beyond. Interaction terms in the model give rise to marginal distributions of population abundances that are non-Poisson even though reproduction of both populations is assumed to be Poisson. The basic assumption about interactions is that any such prey-predator effect over a period of time of residence of two individuals of opposite species can be summarized by a function of current time alone.

INTRODUCTION

Let $\pi = \pi(K)$ denote a K -node network containing at least one absorbing node with no closed proper subsets of nodes. Let S denote a discrete population of individuals that immigrate into π through an initial node i ,

following a prescribed arrival protocol. Once having entered π , individuals never leave; they migrate among nodes, statistically independent of each other, until they enter an absorbing node. If any node is itself composed of a network of subnodes, the assumption is that the conditional residence-time distributions of a typical individual in the node are unique.

Let $(X(t); t > 0)$ denote a continuous-time, finite-state semi-Markov process which describes internode migration of a typical individual upon entering π . $X(t)$ is the node entered by an individual at its most recent change of node prior to time t . $(X(t))$ is characterized by an embedded Markov chain with stochastic, absorbing transition matrix $P = (p_{ij})$, conditional residence-time distribution-function matrix $W = (w_{ij})$, and stochastic interval transition-probability function matrix $F = (f_{ij}(t))$ ($i, j = 1, \dots, K$). The element $f_{ij}(t)$ is the conditional probability that $X(t) = j$, given that the initial state $X(0+) = i$. Assuming time invariance, the conditional probability that $X(t) = j$, given that $X(z) = i$ ($0 < z < t$), is $f_{ij}(t - z)$. Elements of F are functions of elements of P and W [1].

When the immigration process is a Poisson process with nonnegative intensity function $a(t)$, it has been shown [2, 3] that the counts $N_1(t), \dots, N_K(t)$ have the joint probability function

$$P(N_1(t) = n_1, \dots, N_K(t) = n_K) = \prod_{j=1}^K \frac{[m_j(t)]^{n_j}}{n_j!} \exp[-m_j(t)], \quad (1)$$

where $N_j(t)$ is the number of individuals in node j at time t , and

$$m_j(t) = \int_0^t a(z) f_j(t - z) dz \quad (j = 1, \dots, K)$$

Derivation of Equation (1) by use of conditional probabilities is straightforward, resting upon the assumption of a multinomial model of dispersal of individuals upon entering π . That is, the probability of occurrence of the joint event that, upon the arrival of $I = n_1 + \dots + n_K$ individuals initially in node i at time z , there are n_j ($j = 1, \dots, K$) individuals in node j at time t is

$$\frac{I!}{n_1! \dots n_K!} f_1(t - z)^{n_1} \dots f_K(t - z)^{n_K}, \quad (2)$$

where the interval transition probabilities $f_j(x)$ are assumed to be nonvanishing in $(0, t)$.

Other models of dispersal of individuals within π lead to different distributions on the set of possible counts of individuals in nodes. For each combination of arrival process and model of dispersal of individuals upon entering π , a family of parametrized distributions of counts of individuals in nodes is generated. When more than one population is entering π , additional distributions are generated due to assumptions that must be made about the forms of interactions among individuals from different populations. For any two populations simultaneously in π there are four possible forms of interactions: (i) complete independence between populations, (ii) complete two-way dependence, (iii) dependence of one population upon the other, and (iv) the reverse of (iii).

In the present paper we are interested in case (ii) as applied to populations of Canadian lynx and snowshoe hare. Ecological studies of interactions between these populations [4] indicate a weak dependence of hare abundance upon lynx abundance, coupled with a stronger dependence of lynx abundance upon hare abundance. A generalization of Equation (1) to coupled populations of lynx (L) and hare (H) is developed in which regression approximations to interval transition-probability functions describing residence times of the populations in their live states are demonstrated.

Let π denote a four-node network in which nodes 1 and 2 denote live and dead states of hare (H) and nodes 3 and 4 denote live and dead states of lynx (L), respectively. Nodes 2 and 4 are absorbing, while initial nodes 1 and 3 are nonabsorbing. Entry of hare and lynx into their initial nodes 1 and 3 is assumed to be described by corresponding Poisson immigration processes with nonnegative intensities $h(t)$ and $l(t)$. Let the interval transition probability functions for individuals of H , assuming independence of H upon L , be denoted as $f_1(x)$ and $f_2(x)$, where

$f_1(x)$ = conditional probability that a typical individual of H is in the live state x time units after initially entering the live state;

$f_2(x) = 1 - f_1(x)$ = conditional probability that a typical individual of H is in the dead state x time units after initially entering the live state.

Due to the assumption that H is independent of L , the functions $f_1(x)$ and $f_2(x)$ are independent of the parameters of the population L .

Define interval transition probability functions $f_3(x)$ and $f_4(x)$ corresponding to live and dead states of individuals of the population L , independent of the population H . Assuming a multinomial model of dispersal of individuals of H and L upon entering π , the probability of occurrence of the joint event that n_1 and n_2 individuals of H are in the live and dead states at time t , while n_3 and n_4 individuals of L are in the live and dead states at

time t , is

$$P(N_1 = n_1, N_2 = n_2, N_3 = n_3, N_4 = n_4) \\ = \prod_{j=1}^2 \frac{m_j(t)^{n_j}}{n_j!} \exp[-m_j(t)] \prod_{j=3}^4 \frac{m_j(t)^{n_j}}{n_j!} \exp[-m_j(t)], \quad (3)$$

where time t is fixed relative to some arbitrary initial starting time 0, and where

$$m_1(t) = \int_0^t h(z) f_1(t-z) dz,$$

$$m_2(t) = \int_0^t h(z) f_2(t-z) dz,$$

$$m_3(t) = \int_0^t l(z) f_3(t-z) dz,$$

$$m_4(t) = \int_0^t l(z) f_4(t-z) dz.$$

Our purpose in this paper is to drop the assumptions that L is independent of H and H of L , and to demonstrate a regression method of approximating the generalizations of interval transition-probability functions $f_1(x), \dots, f_4(x)$ to the case where their generalizations are functions of parameters of both populations.

A REGRESSION METHOD FOR APPROXIMATING INTERVAL TRANSITION-PROBABILITY FUNCTIONS FOR TWO INTERACTING POPULATIONS

We assume the multinomial model of dispersal of individuals of H and L throughout π , in which dependences of movements of individuals of one population upon individuals of the other are represented in multinomial parameters only. That is, the probability of occurrence of the joint event that upon arrival of $n_1 + n_2$ individuals of H at time z initially in node 1 and arrival of $n_3 + n_4$ individuals of L (independent of H) at time z initially in node 3 there are, $t - z$ time units later, n_1 and n_2 individuals of H in nodes 1 and 2 and n_3 and n_4 individuals of L in nodes 3 and 4 is the (conditional)

multinomial probability

$$\frac{(n_1 + n_2)!}{n_1!n_2!} g_1(t-z)^{n_1} g_2(t-z)^{n_2} \frac{(n_3 + n_4)!}{n_3!n_4!} g_3(t-z)^{n_3} g_4(t-z)^{n_4},$$

where

$$g_1(t) = \alpha_1(t, n_3) f_1(t) \leq 1,$$

$$g_2(t) = \alpha_2(t, n_3) f_2(t) = 1 - \alpha_1(t, n_3) f_1(t),$$

$$g_3(t) = \alpha_3(t, n_1) f_3(t) \leq 1,$$

$$g_4(t) = \alpha_4(t, n_1) f_4(t) = 1 - \alpha_3(t, n_1) f_3(t),$$

and where:

(i) n_1 and n_2 will, in general, be functions of t and of instants in time preceding t .

(ii) α_1 and α_2 are functions of time t , of abundance levels of individuals of L at certain times between 0 and t , and possibly of parameters of the population L . They are not functions of parameters of H or of abundance levels of H at any time.

(iii) An analogous condition holds for α_3 and α_4 .

α_1 and α_2 are not independently determined, since $g_2 = \alpha_2 f_2 = 1 - \alpha_1 f_1$, requiring that $\alpha_2 = (1 - \alpha_1 f_1) f_2^{-1}$. Similarly, α_3 and α_4 are related, implying that α_4 is known as soon as α_3 is determined, given f_3 .

The above assumptions on the separability of g_1 into a product of a function of L alone and a function of H alone, and on the separability of g_3 into a product of a function of H alone and a function of L alone, permit a stepwise method of estimating g_1 and g_3 . First f_1 and f_3 are estimated, assuming no dependences exist between populations, leading to a solution of Equation (3). Then, given f_1 and f_3 , α_1 and α_3 are estimated by regression techniques in which various models incorporating parameters and abundance levels given above are tested. The method yields only approximations, since, in general, the interval transition probabilities $g_1(t)$ and $g_3(t)$ may depend upon much of the previous histories of individuals of one population as they are affected by changes in abundance levels of the other population. Thus, the assumptions given above about prior history of individuals in π as they affect future transitions are no more than approximations to actual but unknown conditions. The solution for the joint probability function analogous

to, and a generalization of, Equation (3) is, unfortunately, not a product of Poisson probabilities as the case is in Equation (3). If the function α_1 upon estimation is not significantly different from unity, the evidence points to independence of H from L . A similar inference holds if α_3 is not significantly different from unity.

If the further assumption is made that counts $N_1(t)$ and $N_2(t)$ of individuals of H in nodes 1 and 2 at time t depend upon counts of individuals of L only in the respective nodes 1 and 2 and only at time t , then the following conditional joint probability holds:

$$\begin{aligned} P(N_1(t) = n_1, N_2(t) = n_2 | N_3(t) = n_3, N_4(t) = n_4) \\ = \prod_{j=1}^2 \frac{[\alpha_j(t, n_3) m_j(t)]^{n_j}}{n_j} \exp[-\alpha_j(t, n_3) m_j(t)]. \end{aligned} \quad (4)$$

If the same assumption is made where H and L are reversed, then

$$\begin{aligned} P(N_3(t) = n_3, N_4(t) = n_4 | N_1(t) = n_1, N_2(t) = n_2) \\ = \prod_{j=3}^4 \frac{[\alpha_j(t, n_1) m_j(t)]^{n_j}}{n_j} \exp[-\alpha_j(t, n_1) m_j(t)]. \end{aligned} \quad (5)$$

The marginal distributions of counts of individuals in nodes 1, 2, 3, and 4 at time t are obtained as solutions to the following sets of simultaneous equations:

$$\begin{aligned} P(N_j(t) = n_j) = \sum_{n_3=0}^{\infty} P(N_3(t) = n_3) P(N_j(t) = n_j | N_3(t) = n_3) \\ (j = 1, 2) \end{aligned} \quad (6)$$

and

$$\begin{aligned} P(N_j(t) = n_j) = \sum_{n_1=0}^{\infty} P(N_1(t) = n_1) P(N_j(t) = n_j | N_1(t) = n_1) \\ (j = 3, 4). \end{aligned} \quad (7)$$

The conditional probabilities in Equations (6) and (7) are obtained directly from Equations (4) and (5) as marginal (Poisson) probability distributions.

TABLE 1
Estimated Sizes of Lynx and Hare Populations^a

Date	Time since date (years)	Population (1000s)	
		Lynx	Hare
1850	0	5	55
1855	5	20	80
1860	10	7	20
1865	15	40	116
1870	20	10	8
1875	25	20	80
1880	30	20	12
1885	35	70	140
1890	40	13	50
1895	45	50	90
1900	50	10	10
1905	55	50	80
1910	60	20	20
1915	65	25	50
1920	70	8	16
1925	75	30	48
1930	80	5	10
1935	85	25	40

^aSource: Sutton and Harmon, *Ecology: Selected Concepts*, Wiley, 1973, p. 195.

For each time t the system (6), (7) must be solved. Numerical solutions are the only feasible means of approximating the marginal distributions. Moments are computed once the marginal distributions are obtained.

CASE STUDY

Records of lynx and snowshoe-hare pelts (Table 1) purchased by the Hudson Bay Company from 1821 through 1934 have been studied by ecologists, statisticians, and population dynamicists [4]. The most important conclusions, summarized by Pielou, are:

- (i) the hare cycle is causal to the lynx cycle;
- (ii) a shortage of plant food rather than an abundance of lynx is the primary cause of decline in snowshoe hare; both hare and lynx abundances are limited by available food;

(iii) both hare and lynx cycles are random, but show autoregressive behavior, indicating internal mechanisms that regulate abundance.

A four-node network as described above was formulated and fitted to data in Table 1. Individuals of H and L populations are assumed to enter nodes 1 and 3, respectively, independently of each other and following Poisson processes. They interact in nodes 1 and 3 (live individuals) as assumed above, creating fluctuations that are in phase (Table 1). An immediate food-limiting effect of declining H is assumed to promote a decline in L . Predation of lynx on hare contributes to a delayed decline in the latter. Peaks in L are followed at five-year intervals by depressions in H , creating a five-year time lag that describes the depressant effect of L on H . It is hypothesized that the real effect of H on L is more pronounced than the real effect of L on H , but that neither interaction effect is dominant as the primary factor accounting for the variation of the time series. According to population dynamicists, quasiperiodic fluctuations in animal abundance occur without interference by other species.

A stochastic compartmental model of hare (H) was formulated and fitted to the data in Table 1 in three steps:

(i) Conceptualize and formalize the model.

(ii) Fit a time-series model of immigration into the live state to the abundance data, assuming an estimated residence-time distribution in the live state, and ignoring any confounding effects of the presence of the lynx population L . Immigration is assumed to be random, i.e., Poisson distributed.

(iii) Fit a model of the conditional expected abundance of L in the live state (node 1), given lynx abundance five years earlier, to H data (Table 1), given the fitted immigration intensity obtained in step (ii) above.

A fourth step, optional, is to adjust coefficients of the conditional-expected-abundance model obtained in step (iii) by linearly regressing the H abundance on the model of conditional expected abundance obtained in (iii) above. The reason for separating steps (ii) and (iii) is the greater influence of non- L factors in creating fluctuations in H abundance, as concluded by population specialists. The order of fitting in steps (ii) and (iii) gains maximum accountability of fluctuations in H due to L -independent effects.

Immigration of individuals of H into node 1 is modeled by a Poisson intensity function $h(t)$ of the form

$$h(t) = A - Bt + C \sin\left(\frac{\pi}{4}(4t - 39)\right),$$

where A , B , and C are estimated from H -data (Table 1). Inclusion of the

term Bt makes not only immigration but net balance in node 1 evolutionary processes, therefore limiting the range of time over which the compartmental model may be validated.

The residence-time probability of a typical individual in node 1 (for short time intervals) is assumed to be proportional to the length of the time interval, leading to a negative-exponential distribution of residence time in node 1. The average time of residence in node 1 for a typical individual was estimated to be three years. Combining the immigration intensity with the residence time density, the mean number of individuals of H resident in node 1 at time t is

$$\begin{aligned} F_1(t) &= \int_0^t h(z) e^{-(t-z)/3} dz \\ &= A[3(1 - e^{-t/3})] - B \left[9 \left(\frac{t}{3} - 1 \right) + 9e^{-t/3} \right] \\ &\quad + C[0.033 \cos \pi t + 0.10\pi \sin \pi t - 0.033e^{-t/3}]. \end{aligned} \quad (8)$$

When $F_1(t)$ was fitted to the H -data in Table 1 by standard multiple regression, the following equation for expected net balance in node 1 was obtained:

$$F_1(t) = 67.6 - 0.42t + 29.4e^{-t/3} - 30.6 \cos \pi t - 92.8\pi \sin \pi t \quad (9)$$

Equation (9) produced a multiple R of 0.84, explaining 70.5 percent of the variability of the H -data.

The next step in model building was to formulate a function $\alpha_1(t, N_3)$ that represents effects of L on H . The function α_1 was hypothesized to have the form

$$\alpha_1(t, N_3) = [1 + N_3(t - 5)]^{B'} e^{-C'N_3(t-5)},$$

where B' and C' are estimated from data. The interval transition probability function

$$g_1(t) = \alpha_1(t, N_3)f_1(t) = [1 + N_3(t - 5)]^{B'} e^{-[C'N_3(t-5) + t/3]}$$

was convolved with $h(t)$ to yield

$$[1 + N_3(t-5)]^{B'} e^{-C' N_3(t-5)} F_1(t),$$

which is the conditional expected net balance of individuals in node 1 at time t , given $N_3(t-5)$. Here $F_1(t)$ is, for each t , a numerical constant, having been fitted earlier to H -data. Since each datum in the H -column of Table 1 is an estimate of the expected net balance of individuals in node 1, the natural logarithm of the ratio

$$\frac{\text{Table 1 value for hare, year } t}{F_1(t)}$$

was set equal to

$$B' \ln[1 + N_3(t-5)] - C' N_3(t-5).$$

A fit to the H -data in Table 1 by standard least squares yielded

$$0.71[1 + N_3(t-5)]^{0.12} e^{-0.003 N_3(t-5)} F_1(t)$$

as the conditional expected net balance of hare in the live state (node 1) in year t .

A final (optional) step consists of regressing hare abundance (Table 1) on the function fitted above, yielding

$$\begin{aligned} E[N_1(t)|N_3(t-5)] = & -2.02 + [1 + N_3(t-5)]^{0.12} e^{-0.003 N_3(t-5)} \\ & \times (54.1 - 0.34t + 23.5e^{-t/3} - 24.5 \cos \pi t - 74.2 \pi \sin \pi t). \end{aligned} \quad (10)$$

Equation (10) explains an additional 2.5 percent of the variability in the H -data in Table 1 beyond that explained by Equation (9). The probability that in year t there are $N_1(t) = n_1$ live hare in the region of Canada represented by the pelt counts shown in Table 1 is, according to the foregoing assumptions leading to Equation (10), Poisson distributed with conditional mean given by the right-hand side of (10), provided the right-hand

TABLE 2
Computed versus Estimated Hare Population^a

Date	Computed hare population	3σ error bounds	Estimated hare population
1850	63	39–87	55
1855	97	67–127	80
1860	35	17–53	20
1865	91	62–120	116
1870	30	14–46	8
1875	89	61–117	80
1880	24	9–39	12
1885	89	61–117	140
1890	20	7–33	50
1895	82	55–109	90
1900	15	3–27	10
1905	76	50–102	80
1910	11	1–21	20
1915	75	49–101	50
1920	6	0–13	16
1925	66	42–90	48
1930	1	0–3	10
1935	59	36–82	40

^aThousands of individuals.

side is positive. To guarantee positive expectation, the final step of fitting should be omitted, yielding the alternative conditional expected net balance given by the previously fitted model. Equation (10) must be further qualified by restricting it to years t for which the expectation remains positive, a restriction required by virtue of the evolutionary character of the hypothesized function $h(t)$.

The conditional Poisson distribution specified by Equation (10) is utilized when solving the sets of equations (6) and (7) for the unconditional distributions of counts of live lynx and hare in year t . Table 2 compares abundance estimates given by Table 1 with estimates computed from Equation (10). Seventy percent of the Table 1 values are within 3σ error bounds of the estimated counts given by Equation (10). Counts $N_2(t)$ of hare represent cumulative numbers of dead individuals whose deaths occurred in the interval $(0, t)$. The stochastic compartmental model predicts a Poisson-distributed number of dead individuals in year t whose mean is the difference between the cumulative number of immigrants to node 1 in the interval $(0, t)$ and the mean number of individuals in node 1 in year t .

Development of a model for the count of lynx (L -data) in the live state (node 3) in year t proceeds in an analogous manner, with certain differences in specification of the functions $l(t)$ and $\alpha_3(t, N_1(t))$. A Poisson immigration intensity for lynx was specified as

$$l(t) = A + B \sin\left(\frac{\pi}{4}(4t - 39)\right),$$

and the function $\alpha_3(t, N_1(t))$ was specified as

$$\alpha_3(t, N_1(t)) = A'[1 + N_1(t)]^{B'},$$

where the coefficients A , A' , B , and B' are to be estimated from L -data in Table 1 following the same procedures as in the case of H -data.

The residence-time distribution of lynx in the live state was assumed to be negative exponential with a mean length of residence of 6 years. Previous theoretical studies of the sensitivity of the shapes of residence-time distributions to expected abundance have shown that shape is much less important than the mean residence time in their effects on expected abundance. No time lag was introduced into α_3 , in view of the data in Table 1 as well as the assumption that reductions in food supplies of lynx have immediate effects upon their abundance.

Following the steps outlined above and demonstrated by construction of the hare model, the following function was fitted to the L -data in Table 1:

$$F_3(t) = 25.1 - 13.2e^{-t/6} - 12.3 \cos \pi t - 232.2 \sin \pi t. \quad (11)$$

Equation (11) yielded a multiple R of 75 percent.

The function $\alpha_3(t, N_1(t))$ was estimated as

$$\alpha_3(t, N_1(t)) = 0.48[1 + N_1(t)]^{0.19}.$$

When the L -data (Table 1) were regressed upon the product $\alpha_3(t, N_1(t))F_3(t)$ and the (optional) fourth step was carried out, the final estimate of the conditional expected net balance of lynx in the live state (node 3) in year t was

$$E[N_3(t)|N_1(t)] = 1.75 + [1 + N_1(t)]^{0.19} \\ \times (10.8 - 5.7e^{-t/6} - 5.3 \cos \pi t - 31.8 \sin \pi t) \quad (12)$$

TABLE 3
Computed versus Estimated Lynx Population^a

Date	Computed lynx population	3σ error bounds	Estimated lynx population
1850	1	0-5	5
1855	33	16-50	20
1860	10	0-19	7
1865	40	21-59	40
1870	10	0-19	10
1875	39	20-57	20
1880	11	1-20	20
1885	43	23-63	70
1890	13	2-24	13
1895	40	21-59	50
1900	10	1-20	10
1905	39	20-58	50
1910	11	1-21	20
1915	36	18-54	25
1920	11	1-21	8
1925	36	18-54	30
1930	10	0-19	5
1935	34	17-52	25

^aThousands of individuals.

The conditional distribution of lynx abundance in node 3 in year t , given the hare abundance in year t , is Poisson distributed with mean given by Equation (12).

An important qualitative difference between Equations (10) and (12) is the nonevolutionary character of (12), a difference readily observed from comparison of the H and L series in Table 1. The multiple R for Equation (12) increases to 81 percent, strongly suggesting the presence of H -dependent effects not accounted for by Equation (11) alone. Comparison of the observed (Table 1) L -abundance with the predicted (Table 3) L -abundance from Equation (12) shows that 84 percent of the data fall within 3σ error bounds of the value computed by Equation (12). The higher percentage of L -data falling within error bounds than that of H -data is explained by the higher regularity of the L -series than that of the H -series, the former being of apparent nonevolutionary character and less pronounced extremes.

Equation (12) and Equation (10) are substituted into the system of Equations (6) and (7). The resulting system, including two normalizing equations, is solved for $P(N_1(t)), \dots, P(N_4(t))$. The system to be solved is finite, as the conditional probabilities obtained from Equations (10) and (12)

diminish in magnitude, so that at some suitable number of terms the series can be truncated. The normalizing equations insure that the sums of terms define proper probability distributions which are not Poisson.

The joint probability function $P(N_1(t), \dots, N_4(t))$, not shown above, is a product involving exponentials but is not Poisson.

DISCUSSION

Equations (10) and (12) may be used to project population abundances beyond the year 1935, although the presence of an evolutionary component in Equation (10) limits the range of its applicability. Assuming the hare do not become extinct, the evolutionary term introduced into $h(t)$ may be modified to approach some constant or slightly fluctuating level. Additional data would be required to estimate the form the modification should assume. Although Equations (10) and (12) produce time-series estimates of the abundance of hare and lynx, the models upon which they are based are quite different from standard time-series models in both concept and development. The models developed above are based upon the theory of Markov-renewal processes and are theoretically capable of providing much detailed information about the movements and disposition of the populations involved, depending in part upon the configuration of the network that describes movements of populations among nodes. The network description of the hare and lynx populations is quite simple, having no loops to permit recycling of individuals among nodes. The complicating factor in the problem analyzed above is the presence of interactions between populations which affect residence times and movements of individuals among nodes. It is a nonlinear modeling problem, which makes it complex in spite of the simplistic nodal structure.

As in most problems of parameter estimation with multiple regression methods, there is more than one procedure that may be followed. Alternative procedures might have been adopted for estimating the coefficients in the H and L models developed above. Stepwise regression might have been used in which steps (ii) and (iii) are combined. Standard multiple regression, combining steps (ii) and (iii), might also have been applied, in which alternative forms of the functions α_1 and α_3 were tested. The main reason for separating steps (ii) and (iii) as done above was to estimate separately the effects of each population upon the other, allowing for maximum contribution of within-population effects to explaining the variability of the time series, consistent with conclusions of population specialists who have studied the L - H interaction problem.

Live births of both hare and lynx are modeled as pure immigration processes, which permit estimates of cumulative births in intervals $(0, t)$

throughout the time interval 1850–1935 and beyond. Immigration intensities $h(t)$ and $l(t)$ model numbers of live births per year, on average, so that mean numbers of birthing events in a period $(0, t)$ are computed by dividing $\int_0^t h(z) dz$ and $\int_0^t l(z) dz$ by the respective mean litter sizes for each population. Representation of reproduction as a random immigration process is not so restrictive as might be thought, as the main consequence of randomizing the rate of arrival of new individuals into the populations is to inflate the variance of the random variables denoting numbers of live births in $(0, t)$ and of the random variables $N_1(t), \dots, N_4(t)$. Since the inflation factor is known for a number of cases involving alternative arrival processes, adjustments to the above variances can be computed, so that other processes of immigration of new individuals into the populations can easily be assumed. In particular, birthing events can be modeled as occurring once or twice per year, with each event producing a random number of live individuals. Means and variances of the net balances $N_1(t), \dots, N_4(t)$ as well as cumulative arrivals in $(0, t)$ have been derived.

Whereas interspecific competitive effects on population balances are represented by inclusion of factors α_1 and α_3 , the effects of intraspecific competition, which are probably significant according to population ecologists, are not represented parametrically as separable factors. Instead, they are approximated through immigration intensities $h(t)$ and $l(t)$ as cyclical components. Inclusion of separable intraspecific competitive factors as model variables introduces additional nonlinearities into the problem and would have carried the modeling problem well beyond the scope of the present investigation.

The residence times of hare and lynx in nodes 1 and 3, respectively, are assumed in the present study to be equivalent to individuals' ages. The residence-time distributions in those nodes are therefore equivalent to the age distributions of the respective populations. The conditional mean times of residence of hare and lynx in nodes 1 and 3, respectively, are

$$\bar{R}_H = \int_0^\infty \alpha_1(t, N_3) e^{-t/3} dt = 3(1 + N_3)^{0.12} e^{-0.003(N_3)} \quad (13)$$

and

$$\bar{R}_L = \int_0^\infty \alpha_3(t, N_1) e^{-t/6} dt = 2.88(1 + N_1)^{0.19}, \quad (14)$$

where the nominal mean life lengths of hare and lynx in the absence of predation are, respectively, 3 and 6 years, and where N_1 and N_3 are assumed to be constants.

TABLE 4
Computed Mean Life Lengths of Hare and Lynx

Lynx population (1000s)	Hare mean life (years)	Hare population (1000s)	Lynx mean life (years)
0	3.00	0	2.88
10	3.88	10	4.54
20	4.07	20	5.13
30	4.14	30	5.53
40	4.15	40	5.83
60	4.10	60	6.29
80	4.00	80	6.64
100	4.00	100	6.92

Table 4 shows that whereas the average age of lynx gradually increases as the abundance of hare increases, the average age of hare increases until the abundance of lynx reaches approximately 39,000, whereupon it begins a steady decline. The relationship between mean lynx age and hare abundance is quite plausible on biological grounds, but the relationship between mean hare and lynx abundance as suggested by Equation (13) is open to question, as it may be an artifact of the data set (Table 1). It may be argued *a priori* that the mean age of hare must decline as predation by lynx increases, i.e., as the abundance of lynx increases. If such a relationship were implicit in the data, then the exponent of $1 + N_3$ ought to have been estimated to be zero or nearly so. According to the data, therefore, it is problematic whether variations in lynx abundance may serve to either increase or decrease the mean age of the population of hare. The exponent estimate of 0.12 is not highly statistically significant.

Nodes 2 and 4 represent the net abundance of hare and lynx, respectively, that are permanently removed from the live, observable states and are not of interest for the present analyses. They are technically sinks in the model and are not required for any of the foregoing observations. They are included without cost for purposes of model completeness. In other studies it may be important to include sinks.

CONCLUSIONS

Introduction of prey-predator effects between two populations introduces interaction terms into the joint probability function describing net balances in nodes, such that marginal distributions must be approximated by solving

simultaneous systems of linear equations in which numerical coefficients are conditional probabilities. By summarizing the cumulative interaction terms between individuals of different populations in terms of functions that depend on the current time only and not on the past, all conditional probabilities are solved as Poisson probability functions. Interaction terms are hypothesized as to form and subsequently estimated by multiple regression techniques, yielding interval transition probabilities that are computed numerically. As is typically the case in population modeling, lack of data prohibits a more adequate testing of the model hypotheses related to interaction effects.

The methodology of representing cumulative interactions between pairs of individuals from different populations in terms of the present is demonstrated to be feasible in terms of both mathematical complexity and parameter estimation. That feasibility is achieved at a cost in realism.

Additional development work on modeling of stochastic interaction effects is needed. Stochastic population models of joint behavior of populations are orders of magnitude more difficult to solve than their ordinary differential- or difference-equation counterparts. There are relationships between stochastic compartmental models and deterministic counterparts, however, that have not been fully exploited in practice. A line of research in stochastic compartmental modeling should be a complete demonstration of these connections.

REFERENCES

- 1 J. Medhi, *Stochastic Processes*, Halstead, 1982, p. 210.
- 2 R. L. Patterson, Continuous time stochastic compartmental models of discrete populations, *Math. Comput. Modelling* 11:975–978 (1988).
- 3 S. M. Ross, *Introduction To Probability Models*, 3rd ed., Academic, 1985, p. 215.
- 4 E. C. Pielou, *Population and Community Ecology*, Gordon & Breach, 1974, p. 92.