A Generalized Model for Birth and Death Mathematical Procedures in the Visitor Management: Evidence from a Protected Area

Ioannis E. Kyritsis and Nikolaos M. Tabakis[*]

Abstract

This paper proposes a generalized model to explain the formation of the annual number of domestic and foreign visitors to Mount Olympus National Park extending the usual birth and death mathematical models. By considering that the time variable takes values from a continuum, we study the dynamics of the population via the calculation of the corresponding marginal changes. The results suggest that these changes could be expressed as the sum of two terms: The first term expresses the change that corresponds to the size of population (phenomenon of "birth" or "death" in the population), while the second term is exclusively a function of time (phenomenon of "out-migration" from or "in-migration" to the population). The proposed method could be used to study similar population structures aiming at the more efficient management of natural areas.

Key Words: *stochastic processes, birth-death models, dynamics of a population*

Introduction

The birth and death stochastic processes constitute a direct generalization of Poisson process and are mainly used by the biological sciences to describe the development of populations (Edelstein-Keshet, 1988; Coquillard and Hill, 1997). In these models it is possible to have a transition from a situation of the system not only to the next (birth) but also to the previous one (death). The realizations of these transitions are not of a constant rate, as in the case of a Poisson process, but they depend on the situation in which the system is found each time, which expresses the size of population.

Moreover, birth-death stochastic processes have been used in queueing systems, since in such a system, the number of customers does not remain constant but is altered depending on arrivals and departures (Kleinrock, 1976; Newell, 1982). In addition, stochastic models that represent real populations or systems subject to various types of catastrophes (e.g. mild environmental changes or extreme natural phenomena, violent human influences) have also been studied extensively (see for example Renshaw, 1991; Matis and Kiffe, 2000; Kyriakidis, 2001). Recently, Economou and Fakinos (2003) presented real applications in population models and queueing systems (stochastic models with catastrophes, general immigration-birth-death models, M/M/1 and M/M/∞ queues). In particular, Kyriakidis (2003) proposed a variation of such a system where the optimal

[*] Correspondence address: Ioannis E. Kyritsis, Department of Economics, Aristotle University of Thessaloniki, P.O. Box 170, 541 24 Thessaloniki, Greece
Tel.: +30-2310-996440, Fax: +30-2310-996426, e-mail: ikyr@econ.auth.gr

policy initiates the controlling action if and only if the state of the process (e.g. number of customers in a queue degree of deterioration of a machine, size of biological population) exceeds a critical number. Such a policy is known as "control-limit policy".

In this paper, the application of a generalized birth-death model in the visitor management of a protected natural area is attempted. It should be noted that, based on the existing sample information, the total number of visitors is divided into subgroups, and for each one of them a different modelling is initially applied.

More specifically, in a recent publication (Parastavrou & Kyritsis, 2001), the structure of the population of visitors to Mount Olympus National Park (Greece) was calculated in relation to the states[1] "new visitor-"0"" and "returning visitor-"1"" for the past two decades. The total number of visitors to the mountain was divided into two strata of non-overlapping populations: **D** (domestic) and **F** (foreigners). The results of the above study are considered in this paper aiming at the appropriate specification, for each one of the strata **D** and **F**, of a model[2] which will be able to describe the formation of the total number of visitors at period *t*. Considering that the time variable *t* takes values from a continuum, we will further study the dynamics of the population via the corresponding marginal changes.

The contribution of the paper is twofold. First, it extends the usual birth and death mathematical models by incorporating a term that is related to the phenomenon of "migration" either from or towards the population under study. Second, the model is applied in a field, which nowadays attracts a growing interest since protected natural areas and the management of the visitors in these areas is considered part of the plan for an integrated sustainable environment.

The remaining parts of the paper are organized as follows: Section 2, presents the data, discusses some methodological issues and reports the empirical results. In section 3, the dynamics of the population via the corresponding marginal changes are studied. Concluding remarks are given in section 4. The relevant tables of the study are presented in the Appendix.

Estimation Results

The Data

The data used in the empirical analysis are obtained from the archives of the Hellenic Federation of Mountaineering and Climbing and are presented in Table 1 (Appendix). More specifically, the table reports the observed total number of visitors to the mountain (*Actual Total*), followed by the observed number of domestic (*Actual Domestic*) and the observed number of foreign (*Actual Foreigners*) visitors. It also reports, in two columns for each one of the strata **D** (Domestics) and **F** (Foreigners), the annual composition of the population of visitors, with respect to each one of the states "new visitor-"0"" and "returning visitor-"1"" (respectively D_0 and F_0, D_1 and F_1).

In order to smooth out the seasonal variability and to determine the long-term trend of the data series, a moving average transformation has been applied: thus, instead of the variable X_t we use $\hat{X}_t = (X_{t-1} + X_t + X_{t+1})/3$: this transformed variable is henceforth used as the visitor variable. The next columns in Table 1, report the estimated values *SmoothD*$_0$ and D_1, *SmoothF*$_0$ and F_1 as derived from the smoothing operation. Using these figures, we estimate models for the annual number of domestic visitors and the annual number of foreign visitors.

A Model for the Number of Domestic Visitors

Given that the annual number of domestic visitors equals the sum of "new" and "returning" visitors, we proceed with separate estimates for each one of its constituent parts as follows:

a. Following Kyritsis *et al*. (1996), the number of "new" visitors (D_0) at period t does not depend directly on the number of visitors in previous periods, but it depends on other factors, such as advertising, information, general conditions, etc. Thus, the model specified for the dependent variable D_0, including time t as the predictor variable, is $D_0=b_1t+b_0+u_t$, where u_t is the disturbance term. We made use of the Microfit 4.1 econometrics package and the initial estimation results were:

$$\hat{D}_0 = 12.326t + 882.870\,. \tag{1}$$

The relevant diagnostic tests appear in Table 2 and reveal the existence of serious autocorrelation. Therefore, we proceeded with the necessary correction (Griffiths *et al*., 1993). The tests for the order of autocorrelation appear at the end of the table, and the model corrected for autocorrelation is:

$$\hat{D}_0 = \underset{(2.797)}{13.393t} + \underset{(38.233)}{861.770}\,, \tag{2}$$

with $R^2 = 0.972$ and $\overline{R}^2 = 0.968$ (the adjusted coefficient of determination), and F=223.254 [0.000]. The numbers in parentheses in (2) are the estimated standard errors for the corresponding coefficients.

b. The number D_1 of "returning" visitors for a period t, by contrast, is supposed to depend on the number of visitors in previous periods. We therefore employ a first-order autoregressive model**.** The initial estimates were:

$$\hat{D}_1 = 0.942D_1(-1) + 152.424\,. \tag{3}$$

On the basis of the relevant diagnostic tests, reported in detail in Table 3, we corrected for autocorrelation and obtained the following estimate of the model:

$$\hat{D}_1 = \underset{(0.047)}{0.912}\, D_1(-1) + \underset{(79.563)}{199.736}\,, \tag{4}$$

with $R^2 = 0.994$, $\overline{R}^2 = 0.993$ and F=1665.1 [0.000].

This autoregressive model can be solved as a difference equation. The solution[3] of the general form of the equation

$$X_{t+1} = \alpha X_t + b \tag{5}$$

is

$$X_t = Ca^t + \frac{b}{1-a}\,. \tag{6}$$

This gives

$$\hat{D}_1 = C \cdot 0.912^t + 2269.727\,, \tag{7}$$

where the parameter C is determined by the initial conditions. An application of the initial condition $D_{1(t=1)} = 862$ to (7) yields $C = -1543.560$, so the solution of (4) becomes

$$\hat{D}_1 = 2269.727 - 1543.560 \cdot 0.912^t\,, \tag{8}$$

and if we take into account that $0.912^t = e^{-0.092t}$, (8) takes finally the form

$$\hat{D}_1 = 2269.727 - 1543.560 \cdot e^{-0.092t}\,. \tag{9}$$

Therefore, in order to predict the total number of domestic visitors, we substitute (2) and (9) in the equation $\hat{D}_t = \hat{D}_0 + \hat{D}_1$, which gives the below final estimate:

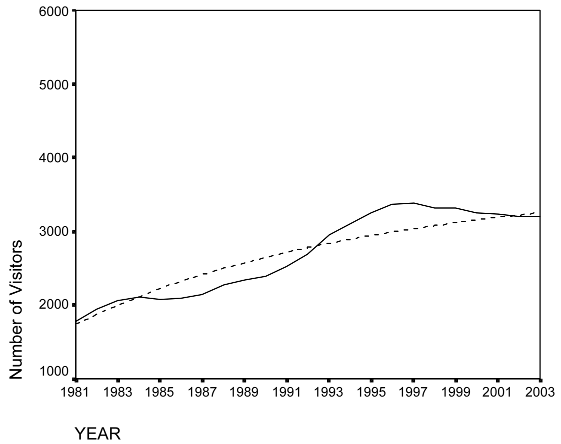
$$\hat{D}_t = 3131.497 + 13.393t - 1543.560 \cdot e^{-0.092t}\,. \tag{10}$$

A Model for the Number of Foreign Visitors

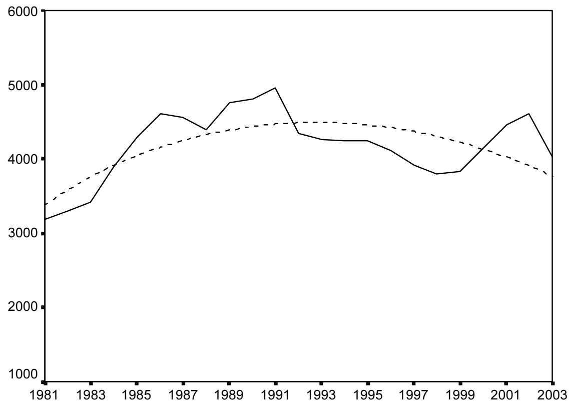
We repeat the same process, using the data for stratum **F** (foreign visitors). However, we found out that a better fit was obtained with a non-linear second-order model, instead of a linear one between the independent variable t and the dependent variable F_0. Thus, after appropriate corrections similar to the ones applied to stratum **D** (which are not reproduced here in order to save space), we obtained the following estimate:

$$\hat{F}_t = (-6.938t^2 + 177.445t + 3359.555) - 233.056 \cdot e^{-0.428t}\,. \tag{11}$$

The fitted values (dotted line) and the observed-smoothed values (solid line) of the total number of domestic and foreign visitors to the region are graphically presented below:



Domestic Visitors



Foreign Visitors

The Dynamics of the Population

Next, if we consider that the time variable t takes values from a continuum instead of taking discrete values t=0,1,2..., that have been used so far for the construction of the equations, we are able to study the dynamics of the population via the corresponding marginal changes.

In stratum **D**, for the measure of the total number of visitors, we use a function of the form

$$D_t = at + b + ce^{kt} . \tag{12}$$

If we take the derivative of (12), we obtain

$$\dot{D}_t = a + cke^{kt} , \tag{13}$$

which presents the rate of change of the visitor population. Finally, if we eliminate the term e^{kt} between the function and its derivative, we obtain the differential equation

$$\dot{D}_t = k \cdot D_t + (a - bk - akt) , \tag{14}$$

and by substituting the numerical values, we obtain:

$$\dot{D}_t = -0.092 \cdot D_t + (301.491 + 1.232t) . \tag{15}$$

Similarly, for stratum **F**, we find that

$$\dot{F}_t = -0.428 \cdot F_t + (-2.969t^2 + 62.070t + 1615.335) . \tag{16}$$

Equations (15) and (16), which express the marginal changes in the populations of strata **D** and **F** respectively, show that these changes equal the sum of two terms. The first term, $-0.092 \cdot D_t$ and $-0.428 \cdot F_t$ respectively, (generally $k \cdot X_t$), depends on the value of X_t of the population and expresses the change that corresponds to its size. This is described as a phenomenon of "birth" or "death", according to whether k>0 or k<0 respectively. More specifically, for the populations under study, we observe that the factor k is negative in both strata, thus expressing the coefficient of death in the population.

The second term, $(301.491 + 1.232t)$ in stratum **D** and $(-2.969t^2 + 62.070t + 1615.335)$ in stratum **F**, depends exclusively on time t and not on the size of the population. This is described as a phenomenon of "migration" from or towards the population, depending, respectively, on whether this second term is negative or positive. More specifically, when the term is negative we refer to "out-migration", while when the term is positive we refer to "in-migration"[4]. For the populations under study, we observe that the term is positive in stratum **D** (assuming no negative values of t), while in stratum **F** the term is positive, when t belongs to the interval [0, 36.01].

Conclusions

In this paper we present a generalized model for the annual number of domestic and foreign visitors to Mount Olympus National Park. The number of "new" visitors at any period t is considered to be independent of the number of visitors in previous periods, while the number of "returning" visitors is considered to be a function of the number of visitors in previous periods. Next, by considering that the time variable t takes values from a continuum, we studied the dynamics of the population via the corresponding marginal changes. The results suggest that these changes are the sum of two terms: The first term $k \cdot X_t$ depends on the value X_t of the population and expresses the change that corresponds to its size (phenomenon of "birth" or "death" in the population). The second term depends exclusively on time t (phenomenon of "out-migration" from or "in-migration" to the population).

The numerical values found for the two strata **D** and **F** can readily be interpreted on the basis of the particular characteristics of the two populations. A very small proportion of the population of domestic visitors who have already visited the Park is lost ($k\approx 9\%$), while the remaining ones visit the area repeatedly. This phenomenon could be explained by the fact that the coefficient of satisfaction of expectations is high, as it was revealed by a relevant survey (Kyritsis *et al*., 1996). By contrast, foreign visitors usually visit just once, but, despite the "deaths" ($k\approx 43\%$), their population is swelled by an in-migration of new individuals, in the same way as it happens with the population of domestic visitors. A further analysis of the different patterns that these models propose with regard to the phenomenon of migration would furnish useful results, while the proposed method could provide more clear inferences and be used to study similar population structures leading to more efficient management of natural areas.

Notes

1 The term "new visitor" is applied to those visiting the area for the first time, and the term "returning visitor" to those who have visited the area at least once in the past.

2 The model can be specified/estimated by applying a number of alterative appropriate techniques, e.g. Box-Jenkins time series analysis, VAR, co-integration. Here, we simply use a heuristic method, which is easy to apply and requires no expertise, making it user-friendly to a broader range of analysts (Makridakis & Wheelwright, 1982) and suitable for the study of changes in the dynamics of a population.

3 For difference equations see e.g. Goldberg, (1961).

4 These are terms that are commonly used in sociology texts.

References

Coquillard, P. and Hill, D. (1997): *Modélisation et Simulation d' Écosystèmes*. Masson, Paris.

Edelstein-Keshet, L. (1988): *Mathematical Models in Biology*. Random House, New York.

Economou, A. and Fakinos, D. (2003): A Continuous-time Markov Chain under the Influence of a Regulating Point Process and Applications in Stochastic Models with Catastrophes. *European Journal of Operational Research,* Volume 149.

Goldberg, S. (1961): *Introduction to Difference Equations.* John Wiley and Sons, New York.

Griffiths, W., Hill, R. and Judge, G. (1993): *Learning and Practicing Econometrics*. John Wiley and Sons, New York.

Kleinrock, L. (1976): *Queueing Systems*, Vol. I & II. John Wiley, New York.

Kyriakidis, E.G. (2001): The Transient Probabilities of a Simple Immigration-Catastrophe Process. *Math. Scientist,* Volume 26.

Kyriakidis, E.G. (2003) Optimal Control of a Simple Immigration Process through the Introduction of a Predator. *Probability in the Engineering and Informational Sciences,* Volume 17.

Kyritsis, I., Papastavrou, A. and Karameris, A. (1996): Survey Interesting Viewpoints of the Visitors of the Mt. Olympus to the Protection, Planning and Development of Area. *Scientific Annals of the Department of Forestry and Natural Environment, Aristotle University of Thessaloniki,* Volume 39/2, (in Greek).

Makridakis, S. and Wheelwright, S. (1982): *The Handbook of Forecasting - A Manager's Guide.* John Wiley and Sons, New York.

Matis, J. H. and Kiffe, T.R. (2000): *Stochastic Population Models*. Springer, New York.

Newell, G.F. (1982): *Applications to Queueing Theory*. Chapman and Hall.

Papastavrou, A. and Kyritsis, I. (2001): Population Structure of a Natural Area: A Markov Chains Approach. *Proceedings of 3rd International Conference. Natural Resources Conservation Services,* Denver – Colorado.

Renshaw, E. (1991): *Modelling Biological Populations in Space and Time.* Cambridge Studies in Mathematical Biology. Cambridge University Press.

APPENDIX

Table 1. Visitors to Olympus National Park during the last twenty five years

YEAR	t	Actual Total	Actual Domestic	Actual Foreigners	D_0	D_1	Smooth D_0	Smooth D_1	F_0	F_1	Smooth F_0	Smooth F_1
1980	0	4749	1673	3076	903	770			2778	298		
1981	1	5266	1746	3520	899	847	914	862	2944	576	2736	446
1982	2	4859	1909	2950	941	968	957	987	2485	465	2778	521
1983	3	5602	2177	3425	1031	1146	974	1082	2904	521	2893	519
1984	4	5945	2084	3861	951	1133	963	1143	3289	572	3324	578
1985	5	6475	2055	4420	906	1149	918	1161	3779	641	3675	624
1986	6	6717	2100	4617	898	1202	896	1197	3957	660	3953	659
1987	7	6925	2124	4801	883	1241	892	1253	4124	677	3913	644
1988	8	6467	2213	4254	897	1316	920	1351	3659	595	3783	616
1989	9	6615	2474	4141	979	1495	930	1418	3566	575	4100	660
1990	10	8241	2356	5885	913	1443	927	1462	5073	812	4144	663
1991	11	6735	2339	4396	889	1450	959	1564	3792	604	4281	682
1992	12	7481	2872	4609	1074	1798	1007	1687	3978	631	3756	596
1993	13	6921	2870	4051	1057	1813	1085	1862	3498	553	3681	582
1994	14	7230	3100	4130	1126	1974	1129	1981	3567	563	3660	577
1995	15	7889	3359	4530	1205	2154	1165	2083	3914	616	3664	577
1996	16	7350	3286	4064	1165	2121	1194	2171	3512	552	3556	559
1997	17	7201	3450	3751	1212	2238	1189	2196	3242	509	3389	532
1998	18	7369	3419	3950	1190	2229	1158	2166	3414	536	3287	516
1999	19	6813	3104	3709	1072	2032	1147	2173	3206	503	3309	519
2000	20	7261	3437	3824	1179	2258	1117	2140	3306	518	3582	562
2001	21	8129	3230	4899	1101	2129	1105	2136	4235	664	3849	604
2002	22	7691	3055	4636	1035	2020	1086	2121	4006	630	3984	626
2003	23	7633	3336	4297	1123	2213	1076	2118	3712	585	3481	548
2004	24	6343	3190	3153	1069	2121			2724	429		

The data of the first five columns derive from the archives of the Hellenic Federation of Mountaineering and Climbing. As visitors were considered all those who, within each year, reached an altitude of 2100 m (where the refuge of the Federation is located) and stayed at least for one night. The rest of the columns are based on the Papastavrou & Kyritsis paper (2001).

Table 2. Estimation Results and Summary Statistics (Dependent Variable: D_0)

Estimation Prior to Autocorrelation Correction			
Regressors	Coefficients	Standard Errors	T-Ratio [Prob]
Constant	882.870	28.561	30.912 [0.000]
T	12.326	2.083	5.918 [0.000]
Diagnostic Tests			
Test Statistic	LM Version	F Version	
A. Serial Correlation	$X^2(1)$=18.101 [0.000]	F(1,20)=73.903 [0.000]	
B. Functional Form	$X^2(1)$=0.292 [0.589]	F(1,20)=0.257 [0.617]	
C. Normality	$X^2(2)$=1.823 [0.402]		
D. Heteroscedasticity	$X^2(1)$=1.878 [0.171]	F(1,21)=1.867 [0.186]	
Estimation Corrected for Autocorrelation			
Regressors	Coefficients	Standard Errors	T-Ratio [Prob]
Constant	861.770	38.233	22.540 [0.000]
T	13.393	2.797	4.788 [0.000]
Parameters of the Autoregressive Error Specification			

$$\hat{u} = \underset{(13.609)}{1.607}\ \hat{u}(-1) - \underset{(-6.979)}{0.824}\ \hat{u}(-2)$$

(T-ratios based on asymptotic standard errors in parentheses)

Log-likelihood ratio test of AR(1) versus OLS: $X^2(1)$=33.807 [0.000]

Log-likelihood ratio test of AR(2) versus AR(1): $X^2(1)$=22.437 [0.000]

Diagnostic Tests:

Table 3. Estimation Results and Summary Statistics (Dependent Variable

Estimation Prior to Autocorrelation Correction			
Regressors	Coefficients	Standard Errors	T-Ratio [Prob]
Constant	152.424	41.253	3.695 [0.0
$D_1(-1)$	0.942	0.024	38.989 [0.0
Diagnostic Tests			
Test Statistic	LM Version	F Version	
A. Serial Correlation	$X^2(1)$=12.319 [0.000]	F(1,19)=24.17 [0.001]	
B. Functional Form	$X^2(1)$=4.783 [0.029]	F(1,19)=5.278 [0.	
C. Normality	$X^2(2)$=1.941 [0.379]		
D. Heteroscedasticity	$X^2(1)$=0.308 [0.579]	F(1,20)=0.284 [0.	
Estimation Corrected for Autocorrelation			
Regressors	Coefficients	Standard Errors	T-Ratio [Prob]
Constant	199.736	79.563	2.510 [0.0
$D_1(-1)$	0.912	0.047	19.558 [0.0
Parameters of the Autoregressive Error Specification			

$$\hat{u} = \underset{(5.208)}{0.743}\ \hat{u}(-1)$$

(T-ratio based on asymptotic standard errors in parentheses)

Log-likelihood ratio test of AR(1) versus OLS: $X^2(1)$=17.416 [0.000]

B. Ramsey's RESET test using the square of the fitted values
C. Test based on skewness and kurtosis of residuals
D. Test based on the regression of squared residuals on squared fitted valu