EigenElephants: When Is Enough, Enough?

David Marks
Jim Sukha
Anand Thakker
North Carolina School of Science and Mathematics
Durham, NC

Advisors: Dot Doyle and Dan Teague

Statement to the Park Management

We develop a system that uses contraceptive darts as the primary method for elephant population control. This method provides a practical alternative to expensive relocation and unpopular culling. Using a statistical model to simulate the changes in the elephant population from year to year, we determine a darting plan that effectively brings the elephant population down to a stable total population of about 11,000, the park's desired target.

Theoretically, this model should accurately predict the structure and size of the elephant population based on the information provided to us about the elephants, such as birthrates, reproductive activity, and life span. Although we had to determine the elephants' survival rates from a rather small sample of data, the survival rates that we determined matched the general information provided. If more accurate survival rates can be found, the model can be adjusted easily by changing a few parameters.

Additionally, we generalize our model to an adaptive darting method that accounts for random fluctuations due to varying survival rates and birth rates, as well as such external influences as immigration, emigration, and poaching. Thus, despite lack of conclusive data, the darting method will effectively control the population even with the random variations introduced by nature.

This method involves the following basic procedure:

- From a survey of the population, determine the approximate population size, age structure, and survival rates. We estimate these from the sample data provided.

The UMAP Journal 21 (3) (2000) 417–430.

- Feed these data into the mathematical model and from it obtain the initial "dosage" (percentage of females to be darted).
  - If relocation of elephants in not a viable option, base dosage for your park is 57%.
  - If it is possible to relocate about 50 to 300 elephants every year, then only 31% of the reproducing female population needs to be darted.

  The females to be darted can be chosen at random, but measures should be taken not to dart the same female twice nor females who are too old or too young to reproduce, as this would reduce the effective proportion of females treated by the contraceptive. Once darting is complete, it is not necessary to track which individuals have been darted, as darting will not be done again until the current dosage wears off, two years later.

- Every two years, count the population and apply the simple formula given in the technical report. We also provide a separate formula for use if the removal of 50 to 300 elephants per year is anticipated.

Under ideal conditions, the park would continue to use the same initial darting plan. However, the population will naturally experience some deviations from the ideal. When surveys show fluctuations in the population, the provided formulas supply the new proportions needed to correct for the deviations.

Our model also tested the survivability of a population after the elimination of a large proportion of elephants. A large natural disaster or widespread disease might cause such a drop in population. Our tests show that when 80% of the population is killed and when survival rates are reduced by 30% for the next 10 years, there is a statistically significant difference between how quickly the population recovers with and without using contraception. However, the darted elephant population still rebounds if darting is stopped, though with a small lag time.

Concern expressed over the validity of the modeling process, especially when the initial data are not completely accurate, is reasonable given the levels of uncertainty that we are working with. However, no matter what method is used for population control, one must have a relatively good idea of the population structure. Our simulations show that our darting plan is flexible and can accommodate variability or inaccuracies in the initial data. This suggests that our model does not depend as heavily on the initial population structure as other methods of population control might. Of course, the best advice we have for increasing confidence in our model is to collect more data. This would provide the most conclusive evidence for the model's accuracy.

Assumptions

- The number of elephants relocated in the past two years is representative of the actual age structure of the current elephant population. One common practice is to relocate entire family units of elephants at once, which would be generally consistent with this assumption.
- The population in the park never differs greatly from the stable state of the population.
- Elephants mate and give birth at a uniform rate throughout the year.
- The population is sufficiently large that we can compute all relevant quantities concerning the population probabilistically.
- The gestation period can be taken to be two years (as opposed to the given twenty-two months).
- The survival rate within ten-year-wide age groups is roughly uniform.

Analysis of the Problem

The nature of this problem suggests that the population should be modeled by a system of difference equations. The data provided by the park are presented in terms of a discrete age distribution. Since the duration of the darts' effectiveness is given in terms of years rather than a fraction thereof, it is appropriate to approach the problem in terms of a discrete time step, namely $\Delta t = 1$ year. This time step sets iterations at one per year and also stratifies the population into cohorts (age groups) of elephants born in the same year.

Given that all elephants die by the age of 70, the problem is reduced to 70 difference equations, one for each age cohort. Such a system is most naturally represented in terms of a matrix equation

$$P_{n+1} = TP_n,$$

where P_{n+1} and P_n are column vectors with 70 rows, in which the ith element represents the number of elephants of age i. The matrix T is 70×70, and each of the elements in the ith row is a coefficient in the ith difference equation. The matrix representation has a powerful advantage over the system of difference equations: T can be manipulated (e.g., by darting) so that it has an eigenvalue of 1, which corresponds to a stable population and age structure.

A matrix A with eigenvalue λ (a scalar) and eigenvector x has the property that $Ax = \lambda x$. For a general population vector P, as $n \to \infty$, we have $A^n P$ approaches x or some scalar multiple of x. The convergence is especially fast if P is initially somewhat similar to x, although small variations of P from x can cause P to converge to a scalar multiple of x instead of to x itself. This relationship suggests the solution to the dilemma of stabilizing the elephant population: If T is manipulated through darting so that it has an eigenvalue 1, then as it is applied to the population of elephants over time, P will converge to the eigenvector; that is, the population will stabilize.

Determining the Transition Matrix

To determine the elements of the matrix equation, consider the structure of the difference equations. The first reduction in the magnitude of the problem is to consider only female elephants. Given that the sex ratio is "very close to 1:1" for adults as well as for newborns, we can consider only females, knowing that the full population can be determined simply by multiplying by two. Hence, the sum of the elements of the P vector should be close to 5,500. The first element of the P vector is the newborn elephants, age 0. The size of this stratum at iteration $i+1$ depends on only the number of reproducing females. The difference equation for the newborn elephants is then

$$(P_0)_{n+1} = \sum_{i=10}^{60} p_i \cdot (P_i)_n,$$

where p_i is the probability that an elephant in the ith age group has a calf that year and $(P_i)_n$ is the ith element of the nth iteration of P; that is, $(P_i)_n$ is the number of elephants in the ith age group in the nth year of iteration. The value of each of the remaining elements in the P vector is determined only by the number of elephants in the previous stratum that survive into that year. This can be written as

$$(P_i)_{n+1} = s_{i-1} \cdot (P_{i-1})_n,$$

where s_i is the probability that an elephant of age i will survive until the next year. This suggests that T is of the form

$$T = \begin{pmatrix} 0 & p_1 & p_2 & 0 \\ s_0 & 0 & 0 & 0 \\ 0 & s_1 & 0 & 0 \\ 0 & 0 & s_2 & 0 \end{pmatrix},$$

but much larger (70×70). Further simplification is desirable.

If P is "close" to an eigenvector, then with each iteration the same numbers of elephants grow into the next age level as they did the previous year; that is, if P is nearly stable, then the population structure should remain relatively constant from year to year. This also means that a larger stratum, say 10 years, has a predictable age distribution. Namely, if a stratum has c elephants growing into it every year with a constant survival rate s over the stratum, then the total number of elephants in the interval is

$$N = c\left(1 + s^1 + s^2 + \cdots + s^n\right),$$

where n is the width of stratum N. The proportion of elephants growing out of stratum into the next stratum is given by

$$\text{Growth} = \frac{cs^n}{c(1+s^1+s^2+\cdots+s^n)} = \frac{s^n(1-s)}{1-s^{n+1}}.$$

Thus, in the steady state, several years of elephants can be grouped together without any loss of information. For the purposes of further discussion, we assume (and verify later) that P is indeed sufficiently close to the eigenvector, and thus we collapse the elephant population into 8 strata, the newborns plus one for each decade up to age 70. The T matrix, now only 8×8, is of a slightly different form, namely,

$$\begin{pmatrix}
0 & 0 & p_2 & p_3 & p_4 & p_5 & p_6 & 0\\
s_0 & s_1(1-g_1) & 0 & 0 & 0 & 0 & 0 & 0\\
0 & s_1g_1 & s_2(1-g_2) & 0 & 0 & 0 & 0 & 0\\
0 & 0 & s_2g_2 & s_3(1-g_3) & 0 & 0 & 0 & 0\\
0 & 0 & 0 & s_3g_3 & s_4(1-g_4) & 0 & 0 & 0\\
0 & 0 & 0 & 0 & s_4g_4 & s_5(1-g_5) & 0 & 0\\
0 & 0 & 0 & 0 & 0 & s_5g_5 & s_6(1-g_6) & 0\\
0 & 0 & 0 & 0 & 0 & 0 & s_6g_6 & s_7
\end{pmatrix},$$

where $g_i = \left[s_i^n(1-s_i)\right] / \left[1-s_i^{n+1}\right]$ is the proportion of elephants that move out of stratum i and into stratum $i+1$ each year. This leaves the determination of survival rates s, probabilities of birth p, and the initial population matrix P to be determined.

Current Age Structure

To determine the survival rate as a function of age, we look to the data provided by the park. From the past two years, we have the sex and approximate ages of the elephants transported out. Under the assumption that elephants were removed fairly uniformly, we take these data to be an accurate representation of the park's overall elephant population. We extrapolate from these data the age distribution of the elephants in the park.

We assume that the elephant population is reasonably stable, in particular, that the overall age distribution of the population is the same for both years. Additionally, since the sex ratio is "very close" to 1:1, we can treat a distribution of one sex as representative of the population's age distribution. We have four samples, one from each sex for each year; we combine these four samples to obtain the relative frequency of elephants at each age (which we scale such that the total number of elephants is 11,000, the park's total population). This distribution is shown in **Figure 1**.

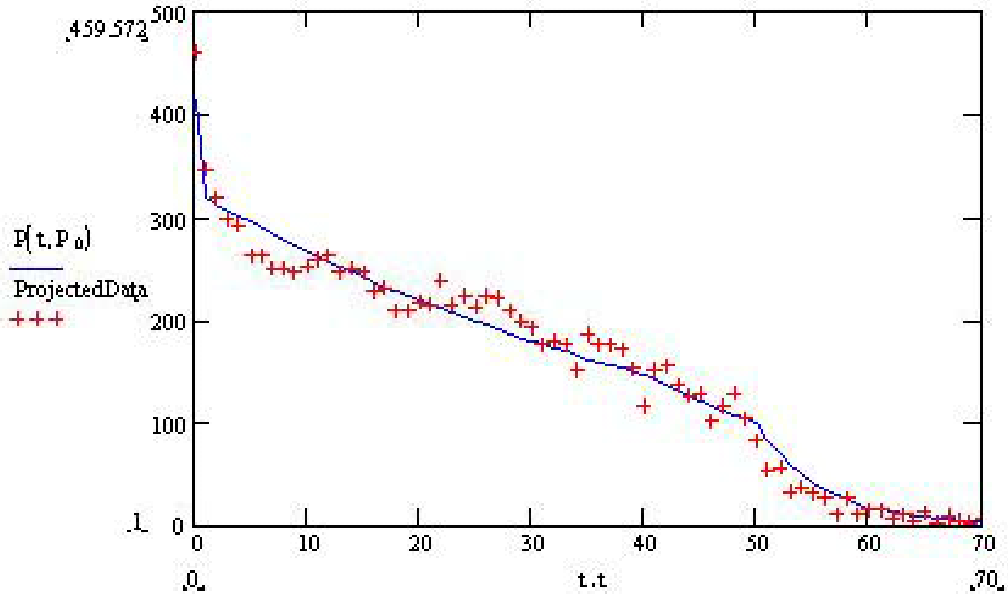

Figure 1. Projected age structure.

Survival Rate

In addition to the age distribution, we are also interested in determining the survival rate of elephants in a given cohort. We begin by following a cohort over time, plotting the number of survivors each year. For a fairly stable population, this relationship is identical to the age distribution. That is, if a population is stable, the age structure is not changing significantly; so as a cohort ages, its size must change to fit the population's age structure.

Since the age structure and cohort survivor data are nearly the same, we use the previously determined age structure to determine the elephants' survival rates as well. Survival rate is defined as the probability that an elephant at a given age survives to the next year. For example, if an elephant has a survival rate of 75% at age 0, the probability that it survives for the next year is 0.75. For large mammals such as elephants, we expect the survival rate in the middle of an elephant's life cycle to remain relatively independent of age while being much lower in very young and very old elephants. Since survival rate governs the change in population from one year to the next and is proportional to the current population, we expect the age structure data to be exponential. The graph in **Figure 1** verifies this prediction.

We plot the natural log of the number of elephants versus age (**Figure 2**) and fit lines through the four major sections of the data, from which we determine the survival rates for four major sections of the population. (**Table 1**).

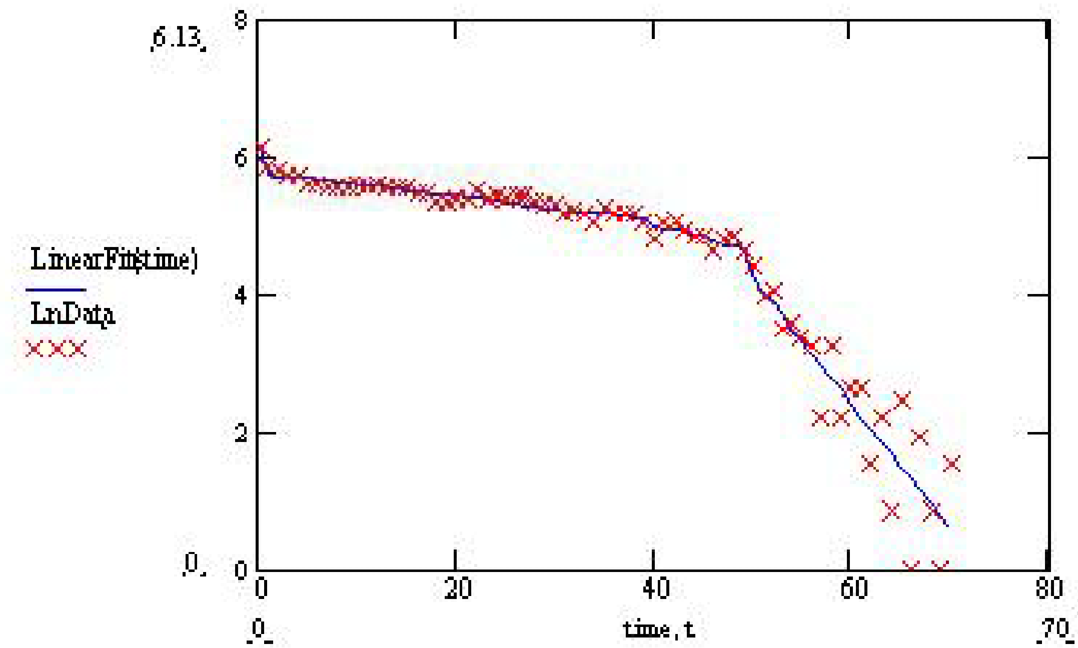

Figure 2. Natural log of elephant data vs. age.

Table 1.
Elephant survival rates.

Age Group (years)	Survival Rate (% / year)
0–1	75
1–50	98
51–60	96
61–70	82

Probabilities of Birth

The information provided by the park suggests that the reproductive rate is constant over all reproducing age groups except the teenage group, where not all of the elephants are reproducing. On average, a female cow produces a calf every 3.5 years with twins occurring 1.35% of the time, indicating that the probability of any given female producing a female calf in a given year is $0.5(1.0135)/3.5 \approx 0.145$. For the teenage group, we assume that about one-third begin conceiving when they are 10, another third when they are 11, and the remaining when they are 12. However, since it takes about two years from conception to birth, the elephants do not actual have young until they are 12. Taking all this into account, the p value for the teenage stratum is 0.7 times that of the other groups. This completes the matrix T (before we start to consider darting).

Introducing the Contraceptive Dart

If the park management darts female elephants at random, then we can assume that the same proportions of each reproducing stratum are sterilized for the two-year period. Hence, if the management keeps some proportion $1-q$ of the population sterile at any given year, then the transition matrix is the same as T above except that the first row has a factor of q in front of the probabilities for birth. The parameter q is the proportion of females reproducing, and this is the value that can be altered to allow T to have an eigenvalue of 1. For an eigenvalue of 1, T must have the property that the determinant $|T - I|$ (an eighth-degree polynomial in q) is 0, where I is the identity matrix. Once q is known and the appropriate T matrix is constructed, an eigenvector can be found quickly. This allows for speculation as to the desirable steady-state elephant population.

The Ideal Solution

Using the values for s and p determined above, the proportion q of females that should be kept reproducing to keep the population stable is about 43%, that is, 57% of the females should be on contraceptives. The appropriate eigenvector associated with a population of 5,500, as well as the extrapolated population estimated from the given data, are shown in **Table 2**.

Table 2.
Eigenvector and extrapolated population.

Cohort	0	1–10	11–20	21–30	31–40	41–50	51–60	61–70
Eigenvector	207	1422	1162	950	776	581	221	181
Extrapolated population	222	1396	1185	1077	833	615	144	27

A measure of the difference of the estimated population from the eigenspace is the cosine of the angle between them. Using the dot product $u \cdot v$, the cosine of the angle between the two vectors is

$$\cos\theta = \frac{EV \cdot EP}{\sqrt{EV \cdot EV}\sqrt{EP \cdot EP}}.$$

For the population estimated from the initial population, we get $\theta = 5.3°$, so the initial population is indeed already close to the eigenvector and our approximation by using 8 strata instead of 70 is valid.

Having set q so that T has an eigenvalue of 1, the matrix is left-multiplied on the population vector. It takes two years for the contraceptive plan to begin to work, due to the two-year gestation period. Using the extrapolated population vector from **Table 2** and assuming the two-year lag period has already passed, **Figure 3** shows the model's prediction of the population over 60 years.

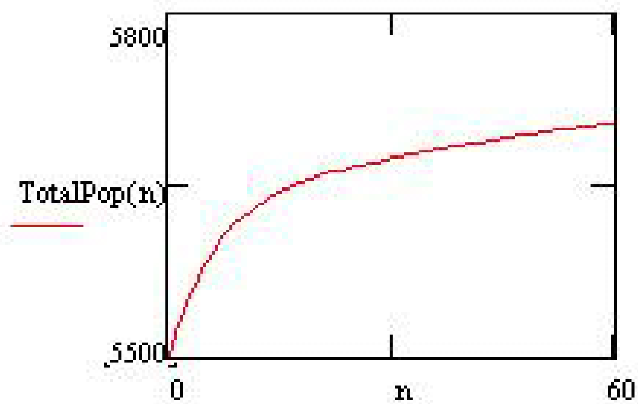

Figure 3. Elephant population over a 60-year period.

The contraceptive rate planned causes a convergence of a population, but it takes a long time. Additionally, the population converges to a higher multiple of the desired eigenvector, with the sum all the female elephants at about 5,700.

The solution to this dilemma is to use not just a single darting plan. Rather, if a more aggressive contraceptive program is used for the first 15 years, the eigenvalue of the matrix will decrease. If the eigenvalue is less than 1, the population eventually begins to drop as it converges to an eigenvector that gets smaller with each iteration (although initially there may be an increase). Once the population has dropped sufficiently, switching to the original darting plan causes convergence to a population distribution with the total number of elephants closer to the desired number (**Figure 4**).

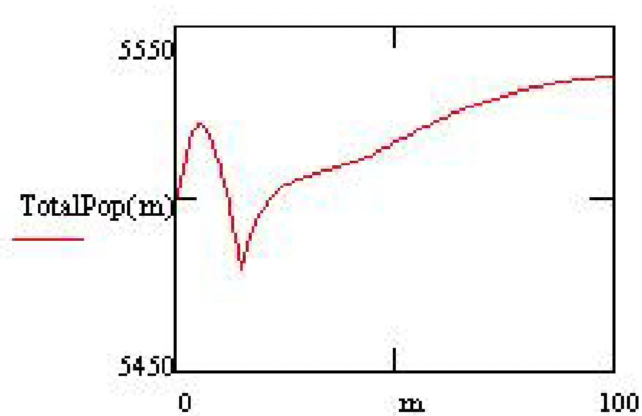

Figure 4. Elephant population over a 60 years with two-phase contraceptive plan.

One particular solution is initially to inoculate about 60% of the population instead of 57%. This small change in the initial 15 years leads to a less variable convergence, where the total elephant population is never more than 100 elephants from the desired 11,000.

The Adaptive Solution

The complication with the two-phase solution is that the exact values for the inoculation are very dependent on properties of the matrix itself, namely, survival rates and birth rate probabilities. Also, more substantial random perturbations in the population, as caused by such phenomena as immigration, emigration, and poaching, can make a static inoculation plan ineffective. Thus, it makes sense to develop a plan that depends on how the population is reacting. If the park can determine the amount of darting needed in the ideal case, then using that value as a base, the park can adjust the actual number of inoculations as required by year-to-year changes in the population.

Given a population at year n, the park would like to know what proportion of females to inoculate. However, changing the proportion of elephants inoculated in year n does not have an effect on the birth rate until year $n+2$, because of the two-year gestation period. In addition, the exact proportion inoculated cannot be adjusted every year independently of the previous year, as the effect of the dart lasts two years.

Thus, there are two possible plans:

- Dart every year, allowing management to raise the levels whenever necessary. If the contraceptive rate every needs to be substantially lower, there will be a one-year lag before a new darting regimen has an effect.
- Dart only every other year, meaning that the management refrains from raising the levels during the off years; but if the dosage needs to go down, there is probability one-half that it can occur immediately.

Both plans have advantages, but the second plan uses substantially fewer darts and will in general be cheaper and require less work to implement.

As the female population changes from 5,500, either due to an eigenvalue not being 1 or to natural perturbations, the number under contraception must be adjusted. The desired proportion $p(N)$ sterile as a function of population N should have the properties that when $N = N_0 = 5{,}500$, $p(N_0) = p_0$ (the value necessary to give the projected T matrix an eigenvalue of 1), and $p(N)$ decreases as N decreases and increases as N increases. To allow for ease of generalization, the amount that p varies should depend on the percentage difference between N and N_0: The effect should be small for small differences in N but should grow quickly enough to constrain N if changes in N are too great. A linear relationship grows too rapidly, suggesting a natural logarithm function of the form

$$p(N) = p_0 + c\,\mathrm{sgn}\left(\frac{N-N_0}{N_0}\right)\cdot \ln\left(\left|\frac{N-N_0}{N_0}\right| + 1\right),$$

where c is a constant that determines how reactive the darting is to changes in the population. We find that values of c ranging from 3 to 5 work well in keeping the population stable (see the **Appendix**). We also find from our simulations that constraining p between some upper and lower bounds increases the stability of the population; a reasonable constraint for p is $0.3 < p < 0.7$.

Contraception with Relocation

If the park managers have the option of removing some elephants in addition to darting, the nature of the problem changes a bit. Consider a population where some number m elephants are to be removed. If they are taken equally from the various strata, then the transition matrix T should be constructed so that it has an eigenvalue $1 + m/N$. This again will allow for a steady state, as each year there will be m more elephants but that many will be relocated. If m is constant, then the problem is solved, as the solution is exactly the same as before, with a slightly different value for p—31%, which is much lower than that amount of contraception use otherwise.

A Disaster

A warranted concern regarding darting is what happens immediately following a natural disaster. We examine the effects by evolving a population for 25 years and then simulating a large natural disaster. This disaster kills 80% of the population and is then followed by a 30% reduction in the survival rates for 10 years. **Figure 5** shows the rebound of a population that begins to reproduce immediately, while **Figure 6** is for a population that must first go through a lag period due to the contraceptives. While the random effects cause some variations depending on the simulation, the overall trend is that the population without contraceptives bounces back faster. Based on 10 simulation runs, the mean population at year 120 with no contraceptive use is 1,246 (SD = 331), while it is only 1,009 (SD = 286) for a group on contraceptives.

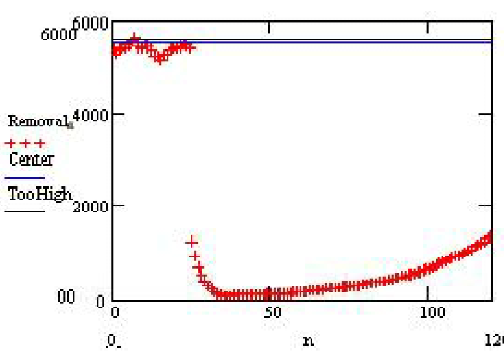

Figure 5. Effect of a disaster on an undarted population.

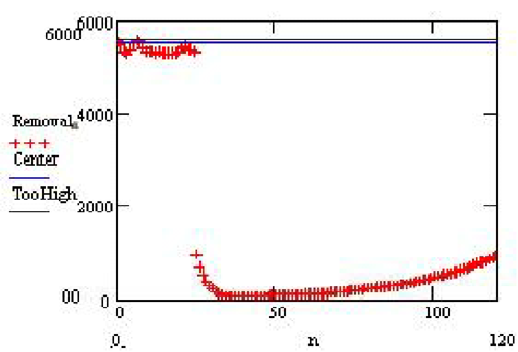

Figure 6. Effect of a disaster on a darted population.

A one-sided t-test of the difference, with df = 17.62, gives a P-value of 0.052, on the border of significance at the 5% level. The opponents of darting may be correct in concerns about an impeded ability of the elephants to grow back. However, the elephant population will still return, if at a slightly retarded rate. Controlling the population without culling elephants seems to justify the risk.

Generalizations

The key to stabilizing a different population at another park is to find the survival rates and birth probabilities, then determine the right value of p that allows T to have an eigenvalue of 1. This process is entirely independent of the actual size of the population, depending on only its age distribution. For small populations, the approximation used to simplify the matrix from 70×70 to 8×8 may begin to break down, but that can be fixed by simply expanding the matrix, which requires just more computer time.

Strengths and Weaknesses

Weaknesses

- Our model for survival rates and age structure depends heavily on the elephant removal data. If these data are not representative of the overall age distribution, then the final population that the model predicts may deviate slightly from the actual value. A more meaningful conclusion on the current age structure cannot be obtained without additional data.
- Our transition matrix considers elephants in 10-year age groups rather than 1-year cohorts. This simplification greatly reduces the size of the transition matrix and allows for quicker calculations, but the approximation may introduce slight inaccuracies. The inaccuracies grow if the population distribution is drastically different from the ideal distribution.
- Elephant populations are discrete quantities, but we approximate them with continuous values. For extremely small populations, this approximation may no longer be valid, especially in the older age groups where there are already very few elephants.
- Our initial model, without adjusting the level of contraceptive darting, is sensitive to changes in survival rates; different values for those can cause the population to converge to a different final value. The modified model that makes adjustments to the level of darting is more capable of handling slight changes in survival rate, but a significant change can still alter the final results.

Strengths

- The final model handles small random fluctuations in the population quite well. These fluctuations add a reality check because they reflect possible error in the park managers' estimate of the population size. The population remains within a reasonable interval around the ideal population, which means the model is not very sensitive to variations in population size.

- The model considers the possibility of some elephants being relocated each year. Relocation when feasible is a preferred method of population control, but the model is not dependent on this possibility.
- Our model can be modified easily to accommodate other parks with different populations and survival rates.

References

World Wide Fund for Nature. 1997. Conserving Africa's elephants—Conservation inside protected areas. `http://www.panda.org/resources/publications/species/elephant/elephant3.html` . Accessed Feb. 5, 2000.

Appendix: Simulation of Adaptive Darting

We determine useful values for c. After each iteration of the simulation, each element in P is multiplied by a random number $\epsilon \sim N(1, 0.01)$. This introduces an effect on the order of 1–2% variation from the predicted value. These effects can be due to errors in the matrix, elephant movements, poaching, or other random effects. We find that effective values for c range from 3 to 5, giving the parks much flexibility in estimating the how much the female population deviates from the desired 5,500.

However, due to the lag involved in the contraceptive's effects, the population meanwhile may deviate farther from 5,500. Placing upper and lower limits on the contraceptive dosage minimizes this problem. Just as with the values for c, the simulation does not change too greatly with different values for the upper and lower bounds on the contraceptive dosage p; a reasonable range seems to be $0.3 < p < 0.7$. The graphs shown in **Figure A1** are eight consecutive random trials, the first four with constraints, the next without. With the exception of the last simulation with constraints, the constraints restrain variability a little but not a lot, suggesting that the park need not worry about exact calculations. Randomness does cause the convergence of the model to disappear. The advantages are that this plan can be started immediately and does not rely on perfectly uniform natural conditions.

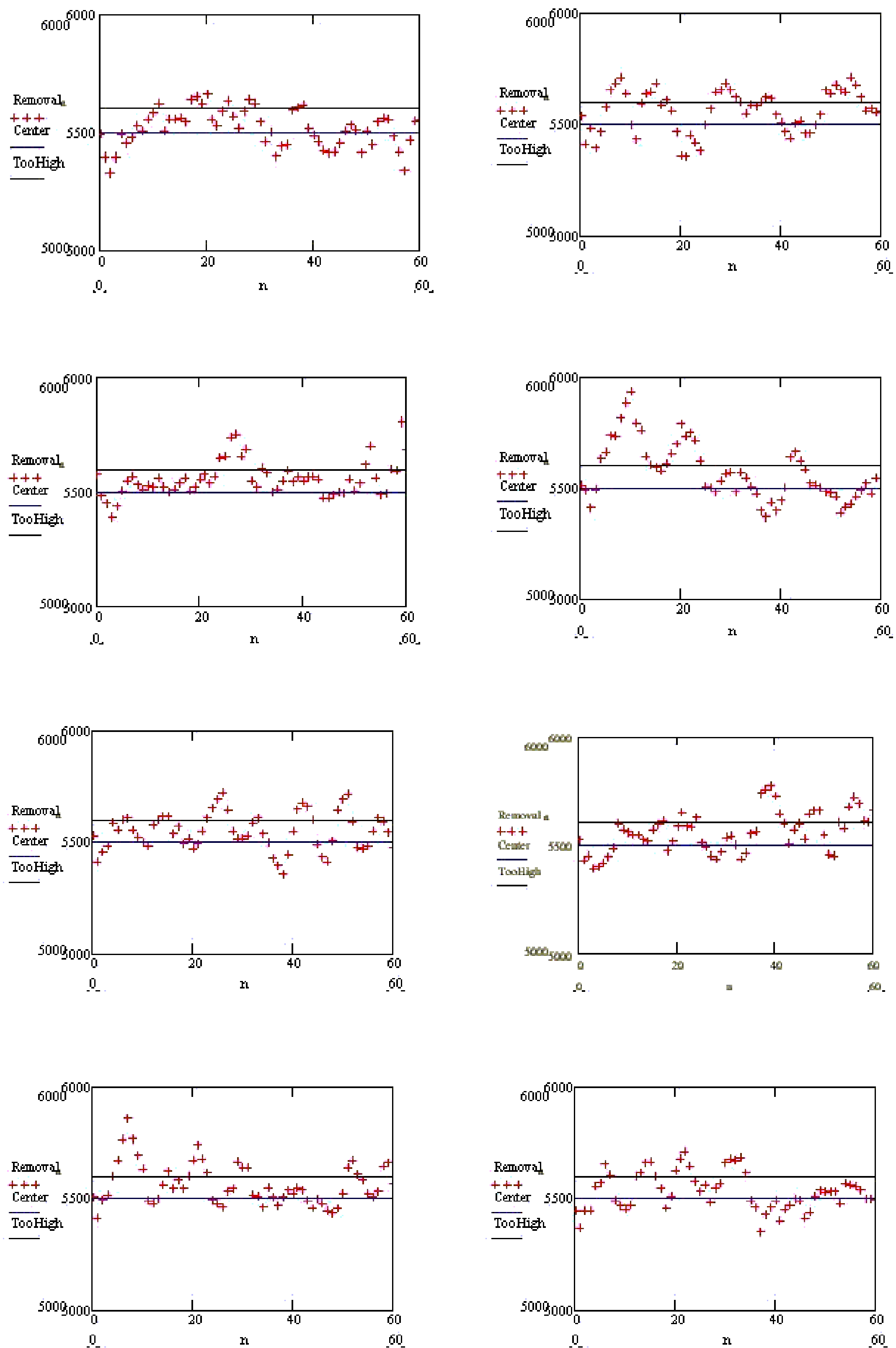

Figure A1. Adaptive contraceptive use with random fluctuations.