An optimal and near-optimal strategy to selecting individuals for transfer in captive breeding programs

S.D. Allen [a,*], Y. Fathi [a], K. Gross [b], M. Mace [c]

[a] Department of Operations Research, College of Engineering, North Carolina State University, P.O. Box 7906, Raleigh, NC 27695, USA
[b] Department of Statistics, College Physical and Mathematical Sciences, North Carolina State University, P.O. Box 7906, Raleigh, NC 27695, USA
[c] San Diego Zoo's Wild Animal Park, 15500 San Pasqual Valley Road, Escondido, CA 92027, USA

ARTICLE INFO

ABSTRACT

As species extinction rates continue to rise, zoos have adopted a more active role in the conservation of endangered species. A central concern is to preserve genetic diversity of zoological populations. Accordingly, when selecting individuals to transfer to new or existing populations, zoo managers must consider the genetic effects on all populations involved. We propose a quadratic integer programming (IP) model to identify a group of individuals to transfer that maximizes genetic diversity within two subpopulations. We then reduce this model to a linear IP formulation and apply it to the California condor (*Gymnogyps californianus*) studbook. After simplifying the linear IP model, optimality is achieved within a reasonable time limit when a limited number of individuals are relocated. We also develop a local improvement algorithm (LIA) to efficiently provide near-optimal solutions when we increase the number of transferred individuals. The LIA quickly obtains optimal solutions when few individuals are transferred and in most cases, the LIA outperforms MetaMK, an existing program used to select animals for transfer.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The combined effect of human population growth, diminished natural habitats, and climatic changes have led to decreases in species diversity and vitality. The International Union for Conservation of Nature (IUCN) predicts that one in four terrestrial vertebrates is currently threatened with extinction (Collen et al., 2008). For critically endangered species, protection provided by zoos may be necessary for long term survival, creating the need for responsible population management strategies and conservation goals.

The preservation of biodiversity relies heavily on the preservation of genetic diversity. Genetic diversity for traits responsive to selection enables adaptive, evolutionary change. Also, in the presence of inbreeding, a loss of genetic diversity is commonly correlated to a rapid reduction in reproductive fitness (Ralls et al., 1988; Crnokrak and Roff, 1999). In populations with structures similar to that of captive, endangered species, retention of the wild population's genetic diversity decreases through generations (disregarding mutations) (Lacy et al., 1995).

It is for these reasons that when creating a new wild or zoological subpopulation, intensive and careful planning is necessary to ensure that all populations under the control of zoo managers remain as genetically unaltered as possible (Frankham et al., 2002). This is a challenge because the addition and removal of individuals change the genetic composition of the population. Frequently, when selecting individuals to remove from one facility and add to another facility (or wild location), the ideal genetic management for one subpopulation is detrimental to the other.

Further, it is usually necessary to address other aspects of management such as the logistic and biological considerations of moving individuals among subpopulations, e.g., transportation expenses, space constraints, and the age distribution of the transferred group. Efficient planning is required to guarantee that the group of individuals selected for transfer best benefits the genetic viability of all subpopulations while satisfying these constraints.

Currently, a computer software program, MetaMK, allows users to manually select individuals for transfer according to various criteria (Ballou, 1999). Possible criteria include sex ratio, the resulting change in genetic diversity among subpopulations, and the relatedness of the individual to others in either subpopulation. This software does not, however, provide transfer recommendations that result in the maximum retention of genetic diversity, nor does it allow for the inclusion of constraints or permit alternative calculations of genetic diversity.

Mathematical optimization is a natural framework for modeling such combinatorial problems subject to various constraints. Published applications of mathematical optimization to managing genetic diversity in captive and reintroduced populations are

* Corresponding author. Present address: Fisheries Ecology Division, Southwest Fisheries Science Center, 110 Shaffer Road, Santa Cruz, CA 95060, USA. Tel.: +1 814 659 2892.
  E-mail addresses: shanae.allen@noaa.gov (S.D. Allen), fathi@eos.ncsu.edu (Y. Fathi), gross@stat.ncsu.edu (K. Gross), mmace@sandiegozoo.org (M. Mace).

somewhat limited (see Kostreva et al., 1999; Vales-Alonso et al., 2003; Fernández and Toro, 1999; Fernández et al., 2001, 2008; Fernández and González-Martínez, 2009; Godoy et al., 2009). Among these, Fernández et al. (2008), Fernández and González-Martínez (2009), and Godoy et al. (2009) address the subject of maximizing genetic diversity when dividing a single population into several subpopulations. Fernández and González-Martínez (2009) suggest minimizing the relatedness of individuals in the subpopulation with the least potential amount of genetic diversity (usually the smallest subpopulation). Godoy et al. (2009) apply this same method when creating new captive populations of Iberian lynx. This technique ensures that the genetic compositions of all subpopulations are as similar as possible. Once all subpopulations are established and running at carrying capacity, Fernández et al. (2008) developed an optimization model and heuristic to simultaneously manage the breeding and relocation of individuals.

We define the Relocation Selection Problem (RSP) as the problem of selecting individuals from zoological subpopulations to transfer to a new zoo or wild location in order to maximize genetic diversity within both subpopulations. Here, we present a quadratic integer programming (IP) model to solve this problem. This quadratic IP model lends itself to several simplifications, thereby reducing its computational complexity. We illustrate the performance of various model formulations by applying data from the California condor (*Gymnogyps californianus*) pedigree. For each formulation, we vary the demand, i.e., the number of transferred individuals and solve the resulting IP model using the well-known IP solver CPLEX 11.

We observe that for relatively small demand values, the execution time of the IP solver is reasonable, but when the demand surpasses ten individuals finding the optimal solution becomes exceedingly time-consuming and perhaps impractical. To remedy this, we design a local improvement algorithm (LIA), a type of heuristic, to solve the RSP. Although the solutions obtained via the LIA are not guaranteed to be optimal, the execution time is greatly improved. Applying data from the California condor studbook to this heuristic results in optimal or relatively near-optimal solutions for the considered range of demand values. We also compare these results with those obtained via the existing heuristic, MetaMK, and observe that for all considered demand values, the LIA achieves comparable or better results.

In the methods section we introduce a few metrics used in population management, present the IP models, and give a short overview of integer programming. Then, in Section 3, we demonstrate the utility of our models by applying data from the California condor studbook. Section 4 includes the design of the local improvement algorithm, its associated computational results, and a comparison to MetaMK. Finally we offer a discussion on model assumptions and possible model extensions.

2. Methods

The following presented models maximize genetic diversity based on information gathered from a complete pedigree. A common estimate of genetic diversity is proportional gene diversity – a measure of genetic variation within a closed population relative to that of the original wild population (the *base population*) (Caballero and Toro, 2000; Falconer and Mackay, 1996; Lacy, 1995). This quantity is denoted GD_t/GD_0, where GD_t and GD_0 are the probability of sampling two distinct alleles at random from the individuals alive at time t and from the base population (assuming an absence of inbreeding), respectively. Analogously, the quantity $1 - GD_t/GD_0$ is termed *average mean kinship* ($\overline{MK}_t$). It is the probability that two randomly sampled alleles from a group of individuals alive at time t are identical by descent, or inherited by the same ancestor (Lacy,

1995). For N diploid individuals alive at time t, the average mean kinship is

$$\overline{MK}_t = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij}}{N^2}, \tag{1}$$

where k_{ij}, the kinship coefficient, is the probability that two randomly chosen alleles, one from individual i and one from individual j, are identical by descent (Ballou and Lacy, 1995; Lacy, 1995; Cockerham, 1969). Hereafter, the subscript t will be dropped for clarity.

Kinship coefficients and average mean kinship may be derived from molecular marker data, or ideally, from a combination of both marker and pedigree information (Toro et al., 1999; Wang, 2001). This information would provide estimates of actual, rather than expected, kinship between individuals.

A caveat to the above definition of average mean kinship is it only applies when there is an even sex ratio among the N considered individuals; in general, this assumption is inappropriate for sexually reproducing species. There exists an alternative calculation of average mean kinship that accounts for uneven sex ratio (Cockerham, 1969). The models that we present here can easily accommodate this alternative calculation, but for ease of presentation and comparison to MetaMK, we use (1) to calculate the average mean kinship regardless of the sex ratio.

Average mean kinship can also be interpreted as the expected mean inbreeding coefficient of N randomly generated offspring if the N considered individuals are hermaphroditic, meaning individuals can self-fertilize and there is no separation of sexes (Cockerham, 1969; Ballou and Foose, 1996; Ballou and Lacy, 1995). Most species of conservation concern are non-hermaphroditic, but a direct relationship between average mean kinship (1) and the expected mean inbreeding coefficient of future generations still exists (Cockerham, 1969). Therefore, minimizing average mean kinship is equivalent to maximizing genetic diversity and minimizing future inbreeding (Lacy, 1995; Ballou and Lacy, 1995).

2.1. Model formulations

Consider a group of N pedigreed individuals (which we will refer to as the *source population*). These individuals may be located among several or hundreds of zoos (and perhaps even wild locations), as long as they are capable of being relocated. Our goal is to remove $d \leqslant N$ individuals from the source population to create a new zoo or wild population (the *transfer population*) in such a way that we minimize the sum of the average mean kinship of both the source and transfer populations. That is, we want to minimize $\overline{MK}_{\text{SOURCE}} + \overline{MK}_{\text{TRANSFER}}$, where $\overline{MK}_{\text{SOURCE}} = \frac{1}{(N-d)^2} \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} (1 - x_i)(1 - x_j)$ and $\overline{MK}_{\text{TRANSFER}} = \frac{1}{d^2} \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} x_i x_j$.

Here, the kinship matrix, K, is an $N \times N$ symmetric matrix that contains kinship values (k_{ij}) for each living pair of individuals, including self kinships. Rows and columns of K are ordered according to the set of studbook numbers, S. The decision variable, x_i, equals one if the individual with studbook number s_i is chosen for transfer from the source population to the transfer site; x_i is equal to zero otherwise. Then, the set of individuals to be transferred can be expressed as a 1-by-N solution vector, **x**.

There are a sizable number of total of possible selection schemes, i.e., $\binom{n}{d}$. This number quickly increases with population size and demand, creating the need for a decision making tool. Therefore, we model this as a quadratic binary integer programming problem where the objective is to minimize the following function,

$$\frac{1}{(N-d)^2} \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij}(1 - x_i)(1 - x_j) + \frac{1}{d^2} \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} x_i x_j \tag{2}$$

This function is subject to two core constraints; (i) the number of selected individuals must satisfy a given demand and (ii) the variables must be binary. Namely,

$$\sum_{i=1}^{N} x_i = d \qquad (i)$$

and

$$x_i \in \{0,1\} \quad \text{for all} \quad i = 1, \ldots, N. \qquad (ii)$$

One comment about this model is that it favors the transfer population over the source population when the number in the source population greatly exceeds number of individuals to transfer (i.e., $1/d^2$ is much greater than $1/(N-d)^2$). While advantageous for small transfer populations, this could result in an undesirable increase in the source population's average mean kinship. One way of preventing this is to simply add a constraint that specifies the maximum allowable average mean kinship of the source population.

While the quadratic IP model may be solved, the nonlinearity of the objective function introduces complications that make computation time prohibitive. This means that finding an optimal solution could potentially take months or even years. On the other hand, solving an integer programming model with a linear objective function is much less computationally intensive due to the ability to use the efficient simplex method in conjunction with the branch-and-bound algorithm (Wolsey, 1998). In brief, the branch-and-bound algorithm separates the optimization problem into smaller subproblems. For each of these subproblems, the simplex method solves the linear relaxation of the IP model (the IP model without the requirement that the solution must be integer). In this way, the simplex method provides a lower bound for the true objective function value (Chvatal, 1983).

An equivalent, linear representation of the quadratic objective function (2) is achieved by introducing a new variable, z_{ij}, such that $z_{ij} = x_i x_j$. This results in the linear objective function

$$\left(\frac{1}{(N-d)^2} + \frac{1}{d^2}\right) \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} z_{ij}$$
$$+ \frac{1}{(N-d)^2} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} - 2 \sum_{i=1}^{N} \sum_{j=1}^{N} k_{ij} x_i\right). \qquad (3)$$

And to avoid introducing the nonlinear constraint, $z_{ij} = x_i x_j$, the following set of linear constraints are substituted:

$$z_{ij} \leqslant x_i \quad \text{for all} \quad i = 1, \ldots, N, \ j = 1, \ldots, N \qquad (iii)$$
$$z_{ij} \leqslant x_j \quad \text{for all} \quad i = 1, \ldots, N, \ j = 1, \ldots, N \qquad (iv)$$
$$z_{ij} \geqslant x_i + x_j - 1 \quad \text{for all} \quad i = 1, \ldots N, \ j = 1, \ldots, N. \qquad (v)$$

After the addition of $\sum_{i=1}^{N} x_i = d$ and $z_{ij} \in \{0,1\}$ for all $i = 1, \ldots, N$, $j = 1, \ldots, N$, a solution to the nonlinear model can be found much more efficiently. Further simplification can be made by observing that without constraints (iii) and (iv) the optimal objective function value remains unchanged. It can be shown that the resulting simplified model has at least one optimal solution that is also optimal for the original problem. This is due to the fact that $k_{ij} \geqslant 0$ which implies that for every optimal solution of the simplified model, $z_{ij} = 0$ when $k_{ij} > 0$ unless constraint (v) forces z_{ij} to equal one. We subsequently compare the computational requirements of these two linear IP models, hereafter referred to as the Linear model (3.1) and Simplified Linear model (3.2), respectively.

While the primary advantage of an IP model is the ability to produce optimal solutions and to assess the nearness of every solution to the lower bound, a limiting factor is that execution time rapidly increases with an increase in demand. This is exemplified by the numerical results presented in the following section.

3. Application

The California condor is one of the world's rarest bird species. Through an unparalleled conservation effort, including comprehensive captive breeding and reintroduction programs, this species is on the way to recovery. Pairwise and self kinship coefficients were obtained from the California condor pedigree as of June 9, 2005 (Mace, 2008), consisting of 150 living individuals. In this illustration, we assume the number of individuals in the source population that are able to be transferred (N) is 150, and the number of individuals in the source population that must be relocated (d) is fixed. Solutions are obtained by using the CPLEX 11.0 Mixed Integer Program solver on a 3.0 GHz AMD Athlon 64 X2 Dual Core processor with 3.4 GB of RAM to run the Linear model (3.1) and the Simplified Linear model (3.2) (ILOG, 2007).

A maximum time limit of two hours was imposed for all problem instances; if the run time reached two hours, the corresponding value reflects the objective function value of the best integer solution found in this time limit. As expected, the Simplified Linear model (3.2) was much more efficient in achieving or coming close to optimality than was the Linear model (3.1) for all considered demand values, d (Table 1). Hereafter, the performance of the Simplified Linear model is used as the basis of comparison for subsequent methods.

When optimality was reached by both models, they provided the same optimal solutions. For instance, when $d = 5$ the optimal objective function value is 0.1515 (Table 1), and the corresponding solution of both models was to select and transfer the group of individuals with studbook numbers 27, 31, 32, 36, and 45. As a result, the average mean kinship of the source population is 0.05151, which is a slight improvement over the average mean kinship prior to transferring individuals, 0.05158. The average mean kinship of the transfer population, 0.10, is the absolute minimum value for a group of five individuals (i.e., all individuals are unrelated and non-inbred). However, as previously stated, the current calculation of $\overline{MK}_{SOURCE}$ and $\overline{MK}_{TRANSFER}$ does not take into account the group sex ratio. Four of the selected individuals are females, which would lead to increased inbreeding coefficients of future generations.

For $d \geqslant 9$, the performance of the Simplified Linear model (3.2) quickly declines relative to lesser demand values, with computation times ranging from 43 min to an excess of 2 h (Table 1). These results illustrate the benefit of a heuristic algorithm – an algorithm that provides solutions as close to the optimum as possible but terminates in a reasonable time limit.

4. Local improvement algorithm

In this section, we introduce a local improvement (search) algorithm (LIA) for solving the RSP and apply this algorithm to

Table 1
Performance of the linear model (3.1) and Simplified Linear IP model (3.2) for varying demand.

	Linear IP model				Simplified Linear IP model			
d[a]	$\overline{MK}_S$[b]	$\overline{MK}_t$[c]	Total[d]	Time	$\overline{MK}_S$	$\overline{MK}_T$	Total	Time
5	0.0515	0.1000	0.1515	5.1 min	0.0515	0.1000	0.1515	1.0 min
6	0.0518	0.0833	0.1351	9.7 min	0.0518	0.0833	0.1351	0.8 min
7	0.0522	0.0714	0.1236	34.1 min	0.0522	0.0714	0.1236	3.2 min
8	0.0527	0.0625	0.1152	93.5 min	0.0527	0.0625	0.1152	6.5 min
9	0.0530	0.0611	0.1141	>120 min	0.0530	0.0586	0.1116	43.0 min
10	0.0530	0.0602	0.1132	>120 min	0.0527	0.0575	0.1102	114 min
11	0.0528	0.0612	0.1140	>120 min	0.0527	0.0558	0.1085	>120 min

[a] Demand.
[b] $\overline{MK}_{SOURCE}$.
[c] $\overline{MK}_{TRANSFER}$.
[d] $\overline{MK}_{SOURCE} + \overline{MK}_{TRANSFER}$.

the California condor pedigree. This algorithm employs the same solution representation, **x**, and the same objective function as the nonlinear IP model (2). A local optimum is found by first randomly generating an initial solution and evaluating the objective function. We refer to this solution as the "current solution." A neighbor to this current solution is then constructed by replacing one transferred individual with an individual that is not currently selected for transfer (Definition I). Therefore, every solution has exactly $d(N − d)$ neighboring solutions. Neighbors are systematically constructed and evaluated until a better solution than the current one is found, at which time this neighboring solution becomes the "current solution." The process is repeated until all neighbors are evaluated without producing a better solution.

Definition I. Given two solution vectors, $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{x}' = (x'_1, \ldots, x'_N)$, where every element x_i and x'_i is either zero or one and $\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x'_i = d$. $\mathbf{x}'$ is said to be a neighbor of $\mathbf{x}$ if $\sum_{i=1}^{N} |x_i − x'_i| = 2$.

The LIA was implemented by using C++ and Microsoft Visual Studio 2008. Results in Tables 2 and 3 reflect the best solution found after repeating this algorithm 100 times, each time starting with a new randomly generated initial solution.

Repetition of the LIA with randomly generated starting points increases the chances of finding a global optimum point. Note that the LIA is strictly a hill-climbing algorithm hence it terminates upon reaching a local optimum point (indeed one or more local optimum points are globally optimal). This could be problematic if the solution space has many local optima. Alternatively we could have used a more robust search method such as the simulated annealing or tabu search that can move out of local optima. We chose the local improvement algorithm (with restarts) because of its relative simplicity and ease of implementation, and for the considered problem it performs reasonably well.

As shown in Table 2, The LIA (with 100 restarts) obtained optimal solutions for all demand values for which optimal solutions were assessed ($d \leqslant 10$). Compared to the performance of the Simplified Linear model (3.2), most solutions were acquired in considerable less time. Subsequently, larger instances of the problem were solved; the solutions and execution times are noted in Table 3. These times ranged from 30 s to 8 min.

Repeating the LIA 1000 times for a particular demand value allows speculation of the topography of the solution space. For $d = 10$, two local minimum values, 0.1123 and 0.1102, were found with a relative frequency of 69.3% and 30.7%, respectively. This amount of repetition leads to higher certainty in finding the global minimum (.1102) but requires a greater elapsed time ($\approx$20 min).

Although the execution times shown in Table 3 are reasonable, the quality of these solutions is not known. In the following subsection we further our computational experiment by solving the

Table 3
Performance of the LIA (100 runs) and MetaMK for varying demand.

d [a]	LIA [b]	Time	MetaMK [b]	Δ [c]
12	0.1066	2.2 min	0.1083	0.0016
14	0.1042	2.4 min	0.1050	0.0009
16	0.1035	3.3 min	0.1035	0.0000
18	0.1023	3.5 min	0.1029	0.0003
20	0.1023	4.2 min	0.1025	0.0002
22	0.1019	5.2 min	0.1026	0.0007
24	0.1019	4.8 min	0.1023	0.0005
26	0.1020	5.1 min	0.1021	0.0001
28	0.1019	7.3 min	0.1021	0.0002
30	0.1019	8.2 min	0.1022	0.0003

[a] demand.
[b] $\overline{MK}_{SOURCE} + \overline{MK}_{TRANSFER}$.
[c] Difference in objective function values.

same collection of instances using a well known heuristic, MetaMK, and compare the results with those obtained via the LIA.

4.1. Comparison to MetaMK

Currently, the two computer software programs used to perform genetic analysis prior to animal relocation are Population Management (PM) 2000 and MetaMK. The former is primarily used to select breeding pairs but it also offers a graphical representation of gene distributions within subpopulations. MetaMK, on the other hand, is specifically designed to select animals for transfer. PM2000 and MetaMK are not fully automated software; they require user input for each iterative selection. Selecting individuals one by one rarely leads to an optimal solution.

MetaMK is an analytical tool that measures the effect of transferring a specific individual on the proportional gene diversity of both the source and transfer populations. For each individual, the change in proportional gene diversity relative to the source and transfer populations is denoted as ΔGD_{SOURCE} and $\Delta GD_{TRANSFER}$, respectively. These values are calculated as the change in proportional gene diversity that would occur if the individual is transferred to the other population and are based on the calculation of average mean kinship according to (1). Individuals in the source population can then be ranked and iteratively chosen for transfer according to the user's preference.

A selection procedure in MetaMK that best coincides with the objective of minimizing the total average mean kinship (or, equivalently, maximizing proportional gene diversity) is to relocate individuals whose removal results in the greatest cumulative change in proportional gene diversity ($\Delta GD_{TRANSFER} + \Delta GD_{SOURCE}$). If two or more individuals result in an equal cumulative change in proportional gene diversity, the individual with the highest mean kinship, mk_i, in the source population is selected. In effect, the individual

Table 2
Performance of the LIA (100 runs) and MetaMK for varying demand.

d [a]	LIA				MetaMK			Δ [e]
	$\overline{MK}_S$ [b]	$\overline{MK}_T$ [c]	Total [d]	Time	$\overline{MK}_S$	$\overline{MK}_T$	Total	
5	0.0515	0.1000	0.1515	0.5 min	0.0517	0.1000	0.1517	0.0002
6	0.0518	0.0833	0.1351	0.5 min	0.0522	0.0833	0.1355	0.0003
7	0.0522	0.0714	0.1236	1.0 min	0.0526	0.0714	0.1240	0.0004
8	0.0527	0.0625	0.1152	0.7 min	0.0523	0.0664	0.1187	0.0036
9	0.0530	0.0586	0.1116	1.0 min	0.0523	0.0617	0.1141	0.0024
10	0.0527	0.0575	0.1102	1.3 min	0.0523	0.0600	0.1123	0.0021
11	0.0527	0.0558	0.1085	2.0 min	0.0524	0.0579	0.1103	0.0018

[a] Demand.
[b] $\overline{MK}_{SOURCE}$.
[c] $\overline{MK}_{TRANSFER}$.
[d] $\overline{MK}_{SOURCE} + \overline{MK}_{TRANSFER}$.
[e] Difference in total values.

with the most redundant amount of genetic information is removed. This method is similar to that used by Ralls and Ballou (2004) when transferring California condor chicks between two wild populations.

In general, the time using MetaMK increases linearly with the number of individuals chosen, although this time would vary with the user and selection criteria. For the considered data set, the average time to select one individual was about seven seconds but it is possible for several individuals to equally satisfy the same criteria, such as $\Delta GD_{\mathrm{TRANSFER}} + \Delta GD_{\mathrm{SOURCE}}$ and mk_i. In this case, trial and error is involved in finding a group to transfer as subsequently selected individuals depend upon those previously chosen.

We used MetaMK to solve for the same collection of demand values that were previously considered using our proposed LIA algorithm. Results are presented in Tables 2 and 3, along with the results obtain via the LIA. In most instances, the LIA produced a better solution than that obtained by MetaMK (i.e., one with a lower objective function value), with a maximum difference of 0.0036 when $d = 8$.

5. Discussion

The task of selecting a group of individuals for transfer that maximizes the gene diversity within two subpopulations can easily be implemented as a linear integer program. For the considered population size, the Simplified Linear model (3.2) obtains optimal solutions in a reasonable amount of time for relatively small demand values. The developed local improvement algorithm provides optimal and near-optimal solutions for a broader range of demand values in minimal computing time. Numerical tests show that the LIA performs as well or better than the alternative heuristic method employed by MetaMK for all considered demand values.

The proposed method for selecting individuals could potentially address two current issues within captive breeding and relocation programs. First, supplying the wild population with more genetically diverse individuals could lead to an increased success rate for reintroduced populations. The success rate of reintroductions is currently very low (13%), although the definition of success is debatable (Fischer and Lindenmayer, 2000). Second, even in zoological populations that allow for control over mate selection, unplanned births are problematic. Cronin et al. (2006) evaluated births and the results of transfers in 35 captive species and found that approximately 50% of births across species resulted from non-recommended pairings. In this setting, ensuring maximal genetic diversity in zoo subpopulations would lessen the chance that unplanned births would decrease the genetic diversity of the population.

There are several direct extensions of the proposed IP models that are applicable to a wider array of management tasks. Demography, especially in reintroduced groups, is extremely important in maximizing reproductive potential. It is quite straightforward to incorporate constraints on sex ratio and age distribution within the transferred group (Allen, 2009). As previously mentioned, the IP models can accommodate alternative calculations of average mean kinship. It is also possible to alter this model to account for preexistent individuals at the transfer site (Allen, 2009). All of these extensions require little additional computational complexity. Finally, to generalize this model further, one could allow for the placement of individuals among two or more transfer sites; however, without further simplification, this would result in a substantial increase in decision variables.

Likewise, the LIA can incorporate the preceding extensions with some manipulation. The usefulness of the LIA depends heavily on the amount of time allotted to make transfer decisions. There are many instances when managers have weeks or months to make a decision. In such cases, the IP model should be attempted but there is no guarantee the IP model will find an optimal solution in the given time limit. Also, new data, i.e., births or deaths may become available and should be included in the model. The LIA could be used to compare the best integer solution to the LIA solution, or to quickly produce a solution that incorporates new information.

In contrast to the iterative method of MetaMK, there are several key advantages to using either an IP model or the LIA in this context. As previously stated, the IP model allows for finding existing optimal solutions and assessing the nearness to the lower bound for every solution, while the LIA can provide at least near-optimal solutions. Also, extensions such as using alternative calculations of average mean kinship or adding additional transfer sites are not possible with MetaMK. Further, the IP model offers much more flexibility when incorporating various selection criteria by adding constraints. For example, one constraint could set a maximum budget for transportation costs, while others could specify individuals that must or must not be transferred together, or could set a limit on the amount of time selected individuals spent in captivity. Quantifying selection criteria in the form of constraints increases consistency and reduces user error.

The relocation of animals, especially for the purpose of reintroduction, is a complex process and the best management practice for one species may not be desirable or feasible for another. The presented models only consider one particular population management goal and ignore many logistic and biological factors. To identify the best strategy, managers must define the program goal based on the biology of the species of interest and the current phase of conservation.

In a broader scope, modeling problems in population management in an optimization framework allows consideration of many factors in a quantitative way, instead of relying on professional judgment that may overlook optimal solutions. Finding optimal management strategies is especially important in conservation programs to best ensure the long term survival of populations.

References

Allen, S.D., 2009. An Integer Programming Approach to Selecting Individuals for Transfer in Pedigreed Populations. NCSU Electronic Theses and Dissertations (etd-12152008-163942) (retrieved 21.01.09).

Ballou, J.D., 1999. MetaMK: Metapopulation Mean Kinship Analysis Software (Version 2.0.0a) [Software]. <http://consgen.mq.edu.au/Links%20and%20Downloads.htm>.

Ballou, J.D., Foose, T.J., 1996. Demographic and genetic management of captive populations. In: Kleiman, D.G., Allen, M.E., Thompson, K.V., Lumpkin, S. (Eds.), Wild Mammals in Captivity. The University of Chicago Press, Chicago and London, pp. 263–283.

Ballou, J.D., Lacy, R.C., 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou, J.D., Gilpin, M., Foose, T.J. (Eds.), Population Management for Survival and Recovery. Columbia University Press, New York, pp. 76–111.

Caballero, A., Toro, M.A., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. Genetics Research 75 (3), 331–343.

Chvatal, V., 1983. Linear Programming. W.H. Freeman and Company, USA.

Cockerham, C.C., 1969. Variance of gene frequencies. Evolution 23, 72–84.

Collen, B., Ram, M., Dewhurst, N., Clausnitzer, V., Kalkman, V., Cumberlidge, N., Baillie, J.E.M., 2008. Broadening the Coverage of Biodiversity Assessments. <http://cmsdata.iucn.org/downloads/broadening_the_coverage_of_biodiversity_assessments.pdf> (retrieved 21.08. 08).

Crnokrak, P., Roff, D.A., 1999. Inbreeding depression in the wild. Heredity 83 (3), 260–270.

Cronin, K.A., Mitchell, M.A., Lonsdorf, E.V., Thompson, S.D., 2006. One year later: evaluation of PMC-recommended births and transfers. Zoo Biology 25 (4), 267–277.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. Prentice Hall, Harlow, England.

Fernández, J., González-Martínez, S.C., 2009. SOFSOG: a suite of programs to avoid inbreeding in plantation designs. Molecular Ecology Resources 10, 393–396.

Fernández, J., Toro, M.A., 1999. The use of mathematical programming to control inbreeding in selection schemes. Journal of Animal Breeding and Genetics 116, 447–466.

Fernández, J., Toro, M.A., Caballero, A., 2001. Practical implementation of optimal management strategies in conservation programmes: a mate selection method. Animal Biodiversity and Conservation 24 (2), 17–24.

Fernández, J., Toro, M.A., Caballero, A., 2008. Management of subdivided populations in conservation programs: development of a Novel Dynamic System. Genetics Conservation 24 (2), 17–24.

Fischer, J., Lindenmayer, D.B., 2000. An assessment of the published results of animal relocations. Biological Conservation 96 (1), 1–11.

Frankham, R., Ballou, J.D., Briscoe, D.A., 2002. Introduction to Conservation Genetics. University Press, Cambridge, UK.

Godoy, J.A., Casas, M., Fernández, J., 2009. Genetic issues in the implementation of Iberian lynx ex situ conservation programme. In: Vargas, A., Breitenmoser, C., Breitenmoser, U. (Eds.), Iberian Lynx Ex situ Conservation: An Interdisciplinary Approach. Fundacin Biodiversidad, Madrid, Spain, pp. 89–99.

ILOG 5.5, CPLEX 11.0, 2007. ILOG [Software]. Sunnyvale, CA. <http://www.ilog.com>.

Kostreva, M.M., Ogryczak, W., Tonkyn, D.W., 1999. Relocation problems arising in conservation biology. Computers and Mathematics with Applications 37, 135–150.

Lacy, R.C., 1995. Clarification of genetic terms and their use in the management of captive populations. Zoo Biology 14, 565–578.

Lacy, R.C., Ballou, J.D., Princée, F., Starfield, A., Thompson, E.A., 1995. Pedigree analysis for population management. In: Ballou, J.D., Gilpin, M., Foose, T.J. (Eds.), Population Management for Survival and Recovery. Columbia University Press, New York, pp. 57–75.

Mace, M., 2008. California Condor studbook. San Diego Zoo, San Diego, CA.

Ralls, K., Ballou, J.D., 2004. Genetic status and management of California condors. The Condor 106 (2), 215–227.

Ralls, K., Ballou, J.D., Templeton, A., 1988. Estimates of lethal equivalents and the cost of inbreeding in mammals. Conservation Biology 2 (2), 185–194.

Toro, M., Silió, L., Rodrigañez, J., Rodriguez, C., Fernández, J., 1999. Optimal use of genetic markers in conservation programmes. Genetics Selection Evolution 31, 255–261.

Vales-Alonso, J., Fernández, J., González-Castaño, F.J., Caballero, A., 2003. A parallel optimization approach for controlling allele diversity in conservation schemes. Mathematical Biosciences 183, 161–173.

Wang, J., 2001. Optimal marker-assisted selection to increase the effective size of small populations. Genetics 157, 867874.

Wolsey, L.A., 1998. Integer Programming. John Wiley and Sons, Inc., New York, USA.