Using historical data to establish baselines for conservation: The black rhinoceros (Diceros bicornis) of the Serengeti as a case study

Kristine L. Metzgera,*, A.R.E. Sinclaira, Kenneth LI. Cambella, Ray Hilbornb, J. Grant C. Hopcraftd, Simon A.R. Mdumae, Robin M. Reichf

a Centre for Biodiversity Research, University of British Columbia, 6270 University Blvd., Vancouver, British Columbia, Canada V6T 1Z4
b School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98185, USA
c Frankfurt Zoological Society, Serengeti National Park, Tanzania
d Serengeti Biodiversity Programme, TAWIRI, Box 661, Arusha, Tanzania
e Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Fort Collins, CO 80523, USA

ARTICLE INFO

Article history:
Received 5 March 2007
Received in revised form
30 May 2007
Accepted 29 June 2007
Available online 21 August 2007

Keywords:
Serengeti Ecosystem
Illegal hunting
Masai Mara National Reserve
Animal offtake
Reintroduction
Spatial model

ABSTRACT

Illegal hunting of rhinoceros in East Africa was widespread in the late 1970s. Today, rhinoceros numbers remain perilously low. The Eastern black rhinoceros (Diceros bicornis michaeli) is restricted to protected areas within Kenya and Tanzania and the few protected areas in Kenya where rhinoceroses are found are reaching carrying capacity. The Serengeti-Mara Ecosystem represents the best potential site for population growth of the Eastern black rhinoceros. Populations within the Serengeti-Mara Ecosystem remain low and augmentation of the current population has been proposed. Using historical census data collected prior to the illegal hunting that occurring in the 1970s, we determined the historical distribution and population of rhinoceros in the Masai Mara National Reserve and Serengeti National Park. The population was approximately 460 animals. We developed a habitat suitability model for the black rhinoceros using the spatial location of historical count data matched with contemporary vegetation and landscape variables. Illegal hunting still remains a significant threat to the rhinoceros. Therefore, we determined areas where the likelihood of a rhinoceros being discovered and then targeted by illegal hunters was highest. This information can be used by managers as a starting point for an assessment for reintroduction if other factors are taken into account. This case study exemplifies the importance and potentially unforeseen applications of long term ecological datasets for species conservation.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

As a result of illegal hunting between 1960 and 1980 black rhinoceros (Diceros bicornis) numbers declined by 95% across their range (from 65,000 to 3800 animals; (Emslie and Brooks, 1999; Muya and Oguge, 2000)). In the 2000s rhinoceros numbers are slowly increasing as a result of intense monitoring and security (Dublin and Wilson, 1998). The Eastern black...
rhinoceros (D. bicornis michaeli), which is a subspecies restricted to Kenya and Tanzania and is considered Critically Endangered by the World Conservation Union IUCN (2004), currently numbers approximately 500 individuals of which 450 are found in protected areas and private conservancies in Kenya (Amin et al., 2006). High growth rates for large herbivores such as rhinoceros can best be achieved if land is not over stocked. Emslie and Brooks (1999) have recommended that populations remain at approximately 75% long term carrying capacity to obtain optimal stocking levels. Due to limited areas available in Kenya, rhinoceros are near or at 75% of the long term carrying capacity in many of these protected areas and conservancies. Therefore, they have recommended that rhinoceroses from populations approaching 75% carrying capacity be translocated to suitable areas within their historic range to augment smaller populations (Emslie and Brooks, 1999). This method would increase their effective size and introduce new genetic material to smaller isolated populations (Emslie and Brooks, 1999).

The Serengeti-Mara Ecosystem which spans Kenya and Tanzania, has been identified as the best potential site for reintroduction to augment the low rhinoceros population that currently lives within its borders (Thirgood et al., 2005). Reintroduction is a widely used tool in recovery plans for threatened and endangered species. The key to a successful recovery plan and reintroduction strategy is to have adequate and reliable biological information of the threatened species (Tear et al., 1995). Between 1977 and 1978, the entire black rhinoceros population within Serengeti-Mara Ecosystem was reduced to only 10 individuals (Borner, 1981; Arcese et al., 1995; Sinclair, 1995). Before this decline little research had been conducted on the black rhinoceros in this ecosystem, and with the speed at which poaching took place, little time was available for data collection. However, we have extracted rhinoceros observations from monthly aerial transect counts that covered the entire Serengeti-Mara Ecosystem collected between 1969 and 1972 (for background see Maddock, 1979). Using these data, we were able to estimate population size and distribution of black rhinoceros prior to 1972 in the Serengeti National Park and Masai Mara National Reserve. We compare our population estimates from these transect counts with an independent total count conducted by members of the Serengeti Research Institute in May 1970 (Sinclair, 1973). While some poaching was observed as early as 1973, we believe that the population was intact during the time of these surveys (A.R.E. Sinclair, pers. obs.).

The World Conservation Union guidelines emphasize the need for an assessment of the availability of suitable habitat as a key component of reintroduction/augmentation planning (IUCN, 1998). Using this historical distribution we assessed the most suitable locations for reintroduction release in the context of habitat suitability. We determined suitable habitat by comparing historical observations with environmental variables. Such environmental variables, including vegetation type, topography and climatic parameters, are primary determinates of broad-scale distribution pattern for large herbivores (Bailey et al., 1996; Fortin et al., 2003). This habitat analysis provides a starting point from which to assess suitable location for animal reintroduction.

To increase a population requires the removal of the cause of the initial decline, illegal hunting in this case (Brambell, 1977; Kleiman, 1989). Therefore, we identified areas within the reserve with the greatest risk of illegal hunting. Illegal hunting of elephants (Loxodonta africana) and rhinoceroses was orchestrated by well organized, funded and armed gangs (Dublin et al., 1990; Sinclair, 1995) and was different from bushmeat hunting. Generally bushmeat hunters were unarmed and relied on rudimentary capture methods such as snaring and poisoned arrows (Wato et al., 2006). Currently, in the 2000s, these organized armed gangs no longer enter the Serengeti-Mara Ecosystem in search of rhinoceroses because rhinoceroses densities are so low that the chance of finding an animal is remote. However, if bushmeat hunters, in the course of their activities, come across a rhinoceros by chance, they kill it themselves or alert rhinoceros hunters of its whereabouts or so that it becomes a target (A.R.E. Sinclair, pers. obs.). Therefore, the risk of a rhinoceros being killed is a function of where illegal bushmeat hunters are most active. The poaching risk model was produced by locating areas within the reserve where poachers were most likely to be present. These areas were identified by combining data on: (i) location of human populations outside the protected area; (ii) accessibility to hunting within the park and (iii) the location of arrests. The density of antelopes can be used as an indicator of human exploitation rates (Setsasas et al., 2007) therefore we used the distribution of the hunting harvest from resident antelopes as an index of illegal hunting activity and hence indirect threat to rhinoceroses. We ascertained hunting offtake by comparing censuses of resident antelopes (using species regularly targeted by bushmeat hunters) from the 2000s with those from the 1990s. The instantaneous rate of change between the two times was an index of the hunting offtake once the natural rate of increase was accounted for. Finally, we integrated the analysis of habitat suitability and poaching risk to determine the most suitable areas within the protected area for rhinoceroses based on both habitat requirements and poaching risk.

Ecological perturbations such as the loss of a species from an ecosystem are unpredictable and can occur rapidly leaving little time for scientists and managers to acquire the necessary baseline data on the animal before it is removed from the system. In the late 1960s the Serengeti Ecological Monitoring Programme started routine aerial surveys over the ecosystem, counting all large species, including the rhinoceros. Less than 10 years later the rhinoceros was almost completely removed by illegal hunting. This study illustrates how long term ecological data can be used for future resource management and conservation objectives.

2. Methods

2.1. Study area

The Serengeti-Mara Ecosystem is situated east of Lake Victoria and northwest of the Ngorongoro highlands and the Rift Valley. The ecosystem covers approximately 27,000 km² (Fig. 1). The Serengeti-Mara Ecosystem is one of the most heavily grazed ecosystems on Earth (McNaughton, 1985).
and is functionally defined by the boundaries of the seasonal migration of over 1.5 million wildebeest (*Connochaetes taurinus*) (Pennycuick, 1975; Norton-Griffiths, 1979). The prevailing northwest direction of storm movements and the location of the Ngorongoro highlands, southeast of the Serengeti National Park produces a rain shadow with reduced precipitation on the southern Serengeti plains at 250 mm yr$^{-1}$ and a gradient of increasing rainfall north-westwards to ~1100 mm yr$^{-1}$.

The Serengeti National Park in Tanzania and the adjacent Masai Mara National Reserve in Kenya are the World Conservation Union category II protected areas; human habitation is restricted to National Park personnel and no hunting or grazing by livestock is allowed within their boundaries. A network of game reserves and conservations areas surround the Serengeti plains (Fig. 1). These areas include Ngorongoro Conservations Area, Loliondo Conservations Area, Ikorongo Game Reserve, Grumeti Game Reserve and Maswa Game Reserve. Game reserves and conservations areas have land use restrictions; licensed hunting but no settlement or cultivation is allowed. Additionally, no hunting is allowed within Ngorongoro Conservations Area, but cultivation (limited to 1 ha plots) is allowed. These game reserves and conservations areas act as a ‘buffer’ zone between Serengeti National Park and areas of unrestricted land use. The Serengeti National Park and Masai Mara National Reserve are protected by anti-poaching activities (funding, infrastructure and park personal are allocated) but some of the game reserves, such as Grumeti Game Reserve, also have extensive anti-poaching activities within their borders. We consider reintroductions for rhinoceros only within the boundary of the World Conservation Union category II protected areas. The short grass plains within Serengeti National Park were excluded from our analysis because rhinoceros did not use this habitat (Frame, 1980). We refer to the combined areas of Serengeti National Park and Masai Mara National Reserve as the ‘protected area’.

We divided the protected area into five zones based on similar zoning by other long term datasets of the protected

---

**Fig. 1 – Map of the Serengeti-Masai Mara Ecosystem in Kenya and Tanzania, East Africa.** The Serengeti National Park and the Masai Mara National Park are protected areas. Game reserves and conservations areas (Ikorongo Game Reserve, Grumeti Game Reserve, Maswa Game Reserve, Ngorongoro and Loliondo Conservation Areas) surround Serengeti. The protected area is divided into zones (North, North Centre, Western Corridor, Southwest, and Short Grass Plains).
area (Sinclair, 1995). We determined population and animal densities within these (Fig. 1). The five zones were:

(ii) North: from the Kenya border south to the Grumeti River.
(iii) North Centre: South from the Grumeti River bounded in the south by the Mbalageti River at Moru and west to Sopa Lodge, bounded in the west from Sopa Lodge to Ikoma gate and in the east by the Serengeti National Park boundary.
(iv) Western Corridor: west of the Sopa Lodge-Ikoma gate line to Ndobaka gate, and bounded in the north by the Serengeti National Park boundary and in the south by a line from Sopa Lodge west to the Serengeti National Park boundary and then to Ndobaka gate.
(v) South West: south of the Mbalageti River at Sopa Lodge, bounded in the southwest by the Serengeti National Park boundary and in the east by a line running south to the boundary near Lake Lagarja.

2.2. Data collection

Distribution and census from transect data (1969–1972): The Serengeti Ecological Monitoring Programme (Norton-Griffiths, 1978) conducted monthly aerial reconnaissance surveys over the Serengeti-Mara Ecosystem (Serengeti National Park, Masai Mara National Reserve, Lolloindo Conservations Area, Ngorongoro Conservations Area, Ikorongo Game Reserve, Grumeti Game Reserve, and Maswa Game Reserve) from August, 1969 to August, 1972. Thirty-one monthly surveys were conducted within this three-year period. Two observers (one on either side of the plane), one recorder and a pilot manned each flight. The plane was flown at 100 m above ground level along flight lines, generally running east to west, which followed Universal Transverse Mercator 10 km demarcations. The observers were not constrained by the distance they could see from the aircraft. At 100 m above ground level rhinoceroses, being similar to the African buffalo (Syncerus caffer) were visible at 250 m from the aircraft based on calibration surveys (Sinclair, 1973). Therefore, the rhinoceroses surveys represented a maximum 500 m swath, which in total was 5% of the ecosystem.

2.2.1. Transect data

Population size was estimated from the 31 reconnaissance surveys (1969–1971). We calculated population size and density (rhinoceroses km^{-2}) of the whole protected area, the Serengeti National Park alone and the five zones within the protected area. We scored numbers of rhinoceros in each transect within the protected area and within each stratum. Using the Jolly method for unequal-size units, we selecting with probability proportional to size (Jolly, 1969). The total population for a single survey was estimated from the sum of the stratum totals and the variance was calculated from the sum of the stratum variances. The population estimates for each of the 31 surveys were then merged using the formulae in Norton-Griffiths (1978) to obtain a weighted mean population over the whole period 1969–1972. The weighting favored those estimates that had the lowest variances.

2.2.2. Total count

In May 1970, an aerial total count of buffalo and elephant was conducted (Sinclair, 1973). Rhinoceroses were also counted during the census. The total count systematically covered blocks of the ecosystem using aircraft that flew a regular path from one end to the other in each block. The western end of the corridor was not surveyed due to aircraft failure.

2.2.3. Visibility bias correction

A visibility bias was used to correct for unseen animals in both the transect and total counts. Not all rhinoceroses were seen from the plane due to animals being hidden in thick vegetation (for description see Goddard, 1967). Goddard compared his aerial counts of rhinoceros from a number of independent aerial counts with a known population in the Olduvai Gorge, close to this ecosystem. At best only 50% of the rhinoceroses were seen from the aerial counts. Similarly, we compared the transect estimate of Masai Mara National Reserve with the known number of 107 rhinoceroses in the Masai Mara National Reserve obtained by Mukinya (1973) using individual recognition. Our transect estimates for Masai Mara National Reserve gave 66 animals, which was 61% of the known number. The total count of 1970 tallied 29 animals in the Masai Mara National Reserve producing a correction factor of 27%. Therefore, we used the correction factor of 61% for the transect data and 27% for the total count to estimate the populations of rhinoceroses in the whole protected area, the Serengeti National Park alone and the separate zones within the protected area.

2.2.4. Transect census data for resident antelopes (1988–2003)

All surveys flown after 1985 were carried out using the same type of systematic reconnaissance flights as those for the early 1970s. Flights were flown at one time during the year (the rainy season) in 1988, 1991, 2001, and 2003 (Campbell and Borner, 1995). We used these surveys to calculate the rate of offtake by humans from resident antelope species. We used counts for kongoni (Alcelaphus buselaphus), impala (Aepyceros melampus) and giraffe (Giraffa camelopardalis) because these three species were widespread and frequently hunted. Flights after 1980 did not include areas in Kenya (Masai Mara National Reserve).

2.3. Independent variables

2.3.1. Precipitation

Precipitation estimates were generated using forty years of monthly rainfall data collected from 58 rain gauges across the ecosystem (Serengeti Ecological Monitoring Program). A computer program, PPTMAP (Coughenour, 2006) was employed to create average monthly and mean annual precipitation estimates for the study region. PPTMAP uses available precipitation data from multiple weather stations, and spatially interpolates the data to develop a grid-cell map of precipitation across the region. The interpolation technique
used was inverse distance weighting, corrected for significant effects of elevation. A regression equation was developed within the program, relating precipitation to elevation, based upon the station data. The slope of the regression line of elevation and precipitation provided a correction of mm rainfall per m elevation difference between any location and any observation station. Precipitation was modeled at 1 × 1 km resolution.

2.3.2. Distance from permanent water sources
At the height of the dry season water sources are found in only a few locations (Sinclair, 1977). Because rhinoceroses need daily access to water (Mukinya, 1977) we measured proximity to permanent water sources. A geographic information system of hydrology was compiled from 55 national (1:50,000) quarter degree topographic maps. Digitzing of the maps was conducted by the Tanzania Wildlife Conservation Monitoring Program in Arusha, Tanzania. All hydrology drainages were identified on the geographical information system and permanent water sources were distinguished from ephemeral rivers and streams by inspection (A.R.E. Sinclair, pers. obs.). The ‘distance to permanent water’ grid-cell map was developed using the ‘Distance Raster’ command in ArcGIS v. 9.

2.3.3. Elevation and site severity index
A geographic information system of contours was constructed in the same manner as that for the hydrology coverage. This geographic information system was converted to a digital elevation model. A digital elevation model is a representation of the topography in grid-cell digital format. From the digital elevation model, we calculated the slope and aspect of each cell. To integrate topographic variables (slope and aspect), we calculated a site severity index for each cell (Nielson and Haney, 1998). The site severity index determines solar radiation and moisture, and may be a predictor of habitat preference for large animals (Nielson and Haney, 1998). Thus,

\[
\text{Site Severity Index} = \sin(\text{aspect} + 225) \times \left(\frac{\text{percent slope}}{45}\right)
\]

where 45 is the maximum percent slope on a site.

2.3.4. Map of vegetation
A geographic information system classification of vegetation cover types was developed using a supervised land cover classification algorithm. Coverage of the area required the use of two separate Landsat 7 enhanced Thematic Mapper scenes (ETM+). These scenes were collected on the same day (2 February, 2000). Extensive field validation verified classification accuracy. Field validation data were either collected or converted using the protocol for classifying East African vegetation described by Grunblatt et al. (1989). The resulting classification produced a grid-cell map of the study area at 90 × 90 m resolution. Independent variables obtained from the vegetation classified map were: (i) the most frequent vegetation class within the defined area (described below); (ii) the variation of vegetation classes within the same area, and (iii) percent canopy cover of shrubs and trees divided into categories.

2.4. Preparation of rhinoceros location data
The transect flight lines were divided into sections 5 km long with a width of 0.5 km. Each rhinoceros observation was assigned a ‘sample location’ in the center of the 5 × 0.5 km block in which it was recorded. We summed all rhinoceros observations from the 31 reconnaissance flights at each sample location to use as the dependent variable. The scale at which the habitat model was applied was limited to the coarsest resolution of the environmental data. This resolution pertains to the precipitation data which were available at a 1 × 1 km resolution. Therefore, in the geographic information system, we defined the sample block as a 1 × 5 km area. This approach was a valid solution to the disparity between the rhinoceros observation sample area (0.5 × 5 km) and the resolution of the precipitation data. The 1 × 5 km block covered the approximate area where a rhinoceros observation was made and represented the general area where a rhinoceros was at the time of the observation. All environmental data (independent variables) were derived from the geographic information system. We used the average of the values of pixels in the surrounding 1 × 5 km area to obtain a representative value of the variable at the center of the block. At each of the sample locations, we extracted the value of the averaged independent variable over the 1 × 5 km block (for example, the average elevation of the 1 × 5 km area).

2.5. Rhinoceros habitat modeling
Modeling of suitable habitat for rhinoceroses was accomplished in two stages. Firstly, using rhinoceros observations from the transect sampling as the dependent variable and elevation and monthly average precipitation as the independent variables, regression analysis was performed using a general linear model (McCullagh and Nelder, 1989). The F-test (\(a = 0.01\)) was used to determine which variables were retained in the general linear model using a step-wise procedure. The rhinoceros observations displayed statistical overdispersion (where observed variance was higher than the variance of a theoretical model), as occurs with many animal observation datasets (Crawley, 2002). Therefore, we specified the Poisson distribution which is more appropriate for abundance data (Jones and Kielland, 2002).

Secondly, we used the fine scale variables such as distance to permanent water and vegetation variables further to explain the residual error produced from the general linear model using a binary regression tree. The regression tree is built through a process known as binary recursive partitioning (Breiman et al., 1984). The regression tree method compares all possible splits among the independent variables using the binary recursive partitioning algorithm. This algorithm maximizes the dissimilarities among groups. Once the algorithm partitions the data into new subsets, new relationships are developed, assessed, and split into new subsets. The algorithm recursively splits the data in each subset until either the subset is homogenous or the subset contains too few observations (e.g. <5) to be split further. The final split is called a terminal node. To avoid over-fitting the model, a 10-fold cross-validation procedure (Efron and Tibshirani, 1993) was used to identify the tree size that minimizes the
Fig. 2a – Observations of black rhinoceros recorded during the surveys of 1969–1972 prior to extensive poaching in 1977–1978. Dots represent the frequency of rhinoceros observations at each particular location. Small dots = 0–3, intermediate = 4–7, and large > 8 observations.

<table>
<thead>
<tr>
<th>Table 1 – Rhinoceros population estimates for the Serengeti National Park and Masai Mara National Reserve and the strata contained within the protected area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transect estimates</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Serengeti National Park and Masai Mara National Reserve</td>
</tr>
<tr>
<td>Serengeti National Park only</td>
</tr>
<tr>
<td>By stratum</td>
</tr>
<tr>
<td>Masai Mara National Reserve</td>
</tr>
<tr>
<td>North</td>
</tr>
<tr>
<td>North Centre</td>
</tr>
<tr>
<td>Western Corridor</td>
</tr>
<tr>
<td>Southwest</td>
</tr>
</tbody>
</table>

a Frame (1980).
b Mukinya (1973).
c Underestimate, entire area was not censuses.
d No estimate available.
total deviance associated with the regression tree. The coefficients generated by the results from the general linear model were used to construct a geographic information system which assigned a value to each grid-cell in the geographic information system of the protected area. The resulting grid-cell map represented the results of the general linear model. The binary recursive partitioning algorithm was then used to generate an additional geographic information system grid of the protected area where each pixel was assigned a value that modeled the error (residual) generated by the general linear model. The final geographic information system surface was the combination of these two models.

We validated the accuracy of the final combined model using a standard cross validation procedure (Efron and Tibshirani, 1993). Details of the accuracy assessment and the cross validation procedure can be found in the appendix.

2.6. Poaching risk assessment and monitoring

Poaching risk is dependent on the probability of rhinoceros encounters with illegal hunters. Because people enter the park for several purposes (e.g. timber, firewood, and meat), we employed two separate analyses to determine the gradient of probability of encountering illegal hunters. The human population to the east of Serengeti National Park and Masai Mara National Reserve (population = 136,000) is primarily comprised of pastoralists who do not consume wild game meat, largely for cultural reasons (Homewood et al., 1987; Bourn and Blench, 1999). In contrast, the agricultural population (1.9 million) that resides on the western border of the park and extends to Lake Victoria derives a proportion of their diet from wildlife consumption (Hofer et al., 2000). Therefore, poaching within the park is largely, if not exclusively, derived from the human population dwelling to the west of the protected area.

2.6.1. Hunters entering the protected area

Firstly, we determined from where hunters were originating and how far within the reserve they traveled. Secondly, we identified where poaching of animals occurred. We used both of these analyses to obtain a map of poaching risk in the protected area. Previous work (Campbell and Hofer, 1995) showed

Fig. 2b – Rhinoceros observations on the total count of May 1970. Dots represent the number of rhinoceros counted at a particular location. Small dots = one rhinoceros, large dots > one rhinoceros. The western end of the Western Corridor was not surveyed due to aircraft failure.
that the number of hunters in the area was a function of the population size and the distance to the reserve boundary. The maximum distance from the home village of an arrested hunter to the park boundary was 45 km. For the first analysis, we calculated the number of hunters in the surrounding area using the current census information (2005) for Tanzania (Bureau of Statistics, Dar es Salaam) and Kenya (International Livestock Research Institute, Nairobi, Kenya, Central Bureau of Statistics, Survey of Kenya 1999). The proportion of hunters per km from the boundary decreased exponentially with distance following:

\[ Y = 0.208e^{(-0.139D)} \quad r^2 = 0.97, \quad P < 0.001 \]  

where \( Y \) is the proportion of hunters per km and \( D \) is distance (km) from the park boundary up to 45 km (Campbell and Hofer, 1995).

The number of hunters per km outside the protected area was calculated by multiplying the proportion of hunters by the census data (humans km\(^{-2}\)). Hunters entering the park were assumed to travel the least distance from their origin to the park boundary. We assigned these hunters to the protected area boundary based on the shortest path from their origin to the boundary. Next, we determined the distance hunters traveled within the protected area by using the location of arrests collected by Serengeti National Park personnel. Rangers recorded the location of both signs of poaching activities (such as snares, deserted poacher camps, sightings of humans) and arrests inside the park. We assumed, as did Hofer et al. (2000), that these indicator data reflected the prevalence of poachers. There were 76 records between June 1991 and February 1992 (Arcese et al., 1995). The number of arrests per km inside the boundary (\( Z \)) was related to distance from the boundary (\( D \)) (\( n = 76, R^2 = 0.67 \)) by:

\[ Z = 9.8305e^{(0.148D)} \]  

From Eq. (2) we determined the distance inside the park that illegal hunters traveled.

2.6.2. Animal offtake rate

In the second analysis, we determined areas of highest animal offtake using resident antelope census data (1988–2003).

---

![Habitat suitability model](image_url)
From sequential censuses we calculated the rate of change of three species populations (impala, kongoni, and giraffe) in each 5 × 5 km grid-cell over the years 1990–2000. These species were widespread and appeared in the hunter kills (Campbell and Hofer, 1995). After accounting for known reproduction and natural mortality (natural yearly rate of increase of 0.10 was used), the remaining mortality was taken to represent human hunting. We averaged the years 1988 and 1991 to estimate the population of the three species for 1990 and averaged years 2001 and 2003 to obtain the 2000 estimate.

In order to combine the two analyses (people entering the park, and animal offtake), we standardized each of the two grid-cell maps by dividing all values in the grid by the highest value of the map. We then added the two maps and standardized the final map by again dividing all cells by the highest value to obtain a scale of 0–1.

2.6.3. Rhinoceros locations for reintroduction based on habitat suitability and illegal hunting risk

Finally, we combined the habitat suitability map and the risk assessment map to identify areas in the reserve that maximize suitable habitat and minimize risk from illegal hunting.

3. Results

3.1. Rhinoceros population estimation

3.1.1. Transect counts

Fig. 2a shows the distribution of rhinoceros sightings. The highest numbers of sightings were in northern Serengeti National Park and Masai Mara National Reserve. The next highest numbers were in the west, and along Olduvai gorge in NCA. Using the 31 monthly surveys in 1969–1972 we calculated a weighted mean total for the Serengeti National Park and Masai Mara National Reserve of 461 rhinoceroses after applying the visibility correction (Table 1). We also obtained estimates for each zone separately. We excluded the short grass plains and areas outside the protected areas.

We estimated that the Serengeti National Park had 336 (±37) rhinoceroses. Because we calculated visibility bias using Masai Mara National Reserve data, the estimate for that area was the same as the individually known number from Mukinya (1973). The highest number of individuals occurred in the north (140 ± 22), and the lowest in the southwest at 53 (±15). Densities in the whole protected area and Serengeti National Park alone were 0.03 animals km$^{-2}$. Density was highest in the North zone (0.08 animals km$^{-2}$) and Masai Mara National Reserve (0.06 animals km$^{-2}$) and lowest in the North Centre and the Southwest (0.02 animals km$^{-2}$). The Western Corridor had only slightly higher densities than the Southwest and North Central zone with 0.03 animals km$^{-2}$.

3.1.2. Total counts

Fig. 2b shows the distribution of rhinoceroses observed during the total count of 1970. The count recorded 140 rhinoceros, of which 119 were observed within the protected area (Table 1). The North zone had the largest population at 185 animals and the north centre and the southwest had the lowest population (37 and 33, respectively).

3.2. Rhinoceros habitat modeling

Biotic and abiotic independent variables were used to predict the rhinoceros transect observations using the general linear model and the regression tree. We combined the output from the general linear model and the regression tree to produce a grid-cell map of preferred habitat by rhinoceroses (Fig. 3). The combined general linear model and regression tree explained 64% of the variability in rhinoceros observations. Elevation and precipitation (March, June, November and December) were important in explaining variability in rhinoceros locations using the general linear model. Rhinoceros were slightly negatively correlated with elevation, positively correlated with March and June precipitation, and negatively correlated with December and November precipitation. Intermediate elevation zones (1300–1600 m) overlapping with high precipitation at key months of the year (June and March), predicted where suitable habitat was located.

| Table 2 – Variables used in the final general linear model and regression tree |
|-----------------------------|-----------------------------|-----------------------------|
| Independent variables       | General linear model         | Regression tree             |
| Abiotic variables           | Site severity index          | *                          |
| Elevation                   | Variation in elevation      | *                          |
| Distance to permanent water | *                          | *                          |
| Precipitation               |                            |                            |
| January                     |                            |                            |
| February                    |                            |                            |
| March                       |                            | *                          |
| April                       |                            | *                          |
| May                         |                            | *                          |
| June                        |                            | *                          |
| July                        |                            | *                          |
| August                      |                            |                            |
| September                   |                            |                            |
| October                     |                            |                            |
| November                    |                            |                            |
| December                    |                            |                            |
| Mean annual precipitation   | *                          |                            |
| Biotic variables            | Variation in vegetation type| *                          |
| Vegetation type             | Percent cover shrubs        |                            |
| Variation in percent cover shrubs | Percent cover trees     |                            |
| Variation in percent cover trees | *                     |                            |

Variables used in the regression trees describe the errors associated with the results from the general linear model. Combined general linear model and regression tree model results can be seen in Fig. 4.
The regression tree explained the residuals produced from the general linear model using a combination of biotic and abiotic variables. The independent variables used in the regression tree included dominant vegetation type, variation in percent cover of trees, mean annual precipitation, site severity index, elevation, variation in elevation, and distance to permanent water sources. Grassland and woodland in elevation zones below 1600 m contributed to explaining the fine-scale variability of the general linear model. Distance from permanent water was important for determining rhinoceros habitat in certain vegetation types. Suitable habitat was located at a distance of 2 km from permanent water sources in scrubland and woodland and within 15 km in all other vegetation types. In all vegetation types and elevation zones, distances greater than 15 km from permanent water sources did not help to explain fine-scale variability of the general linear model. A list of variables included in the general linear model and the final regression tree are given in Table 2. The general linear model explained 29% ($R^2 = 0.29$, $p$-value < 0.0001) of the variation in rhinoceros locations while the regression tree explained an additional 35%. Cross validation indicated that prediction bias for the model was nominal and the estimated uncertainty of the model was statistically consistent with the true errors. Full model evaluation and results using the cross validation can be found in the appendix.

### 3.3. Distribution of illegal hunting

#### 3.3.1. Hunters entering the protected area

There are three major areas of high human densities (250–450 people km$^{-2}$) either in close proximity to or bordering the Serengeti National Park boundary, which indicates where hunters occur (Fig. 4a). High densities of hunters are located along the edge of the Maswa Game Reserve, Grumeti Game Reserve and Ikorongo Game Reserve (Mugumu) (Fig. 4a).

How far people live from the protected area largely influenced whether they hunted illegally for wildlife within the protected area (Campbell and Hofer, 1995). Ikorongo Game Reserve and Maswa Game Reserve acted as buffers to the Serengeti National Park by increasing the distance hunters needed to travel to gain access to the park from origins.

![Fig. 4a – Hunter density to the west of the protected area. High hunter densities are represented by high human populations in close vicinity to a protected area border. Scale is hunters per km$^{-2}$.](image)
outside of these controlled areas. To a lesser extent, Gru-
meti Game Reserve also served as a buffer but the narrow
width of the Grumeti Game Reserve (widest point is 8 km)
allowed hunters to travel through it to enter the park
(Fig. 4b).

There was an exponential decline in hunters arrested with
distance inside the protected area, indicating that they re-
mained close to the boundary. Hunters that entered the pro-
tected area from locations along the border, that are not
adjacent to a buffer, traveled the furthest (Fig. 4b). Due to
the geographical shape of the Western Corridor, the majority
of community land within this region is within 20 km of the
protected area boundary making it easily accessible on foot.
West of the northern Serengeti, close to the Kenya–Tanzania
border, there is another heavily populated area associated
with the towns of Mugumu and Tarime. The absence of a buf-
fer zone in this area allowed hunters to travel further into the
protected area (Fig. 4b).

3.3.2. Animal offtake rates
Animal offtake rates varied between ~0.5 animals km$^{-2}$ yr$^{-1}$
in the eastern Serengeti to a high of 1.6 animals km$^{-2}$ yr$^{-1}$
in the Western Corridor (Fig. 5a). Offtake rates were highest in
the North, North Centre and Western Corridor, and decreased
with distance from the western Serengeti National Park
boundary towards the east, particularly in the center of Ser-
engeti National Park along the Orangi River.

The values for hunter presence (Fig. 4b) and animal offtake
(Fig. 5a) were combined and again standardized to the maxi-
mum cell value of 1, to produce a surface for poaching risk
(Fig. 5b). Areas of high poaching risk were located in the Wes-
tern Corridor and the far northwest of the protected area.
Although the Ikorongo Game Reserve reduced hunter infiltra-
tion into the protected area, hunting occurred to the east of
this reserve. The Grumeti Game Reserve deterred some hunt-
ers from crossing into the Western Corridor compared to its
southern edge which experienced the highest poaching

Fig. 4b – Hunters entering the protected area. Contours are based on the frequency of location of hunters arrested within the
protected area. The highest probability of arrests lies closest to the western boundary and decreases exponentially eastwards
as indicated by the distance between contours. High density of contour lines signifies a higher probability of the presence of
hunters within the protected area.
pressure in the park. Thus, there was a general west to east decrease in poaching risk.

3.4. Areas for reintroduction

We combined the surfaces for poaching risk and habitat suitability and then standardized to unity (dividing the map by the highest value on the map), to obtain a surface for reintroduction suitability (Fig. 6). Although the northwest had the highest rhinoceros densities (Figs. 2a, 2b, Table 1) and the best habitat (Fig. 3), it also had high poaching risk (Fig. 5b), which reduced its overall suitability for reintroduction. Despite this poaching risk, the north proved to be the best location, especially further away from the western edge. In contrast, the Western Corridor had moderate rhinoceros densities and habitat suitability, and these were greatly overshadowed by high poaching risk. Thus, the western half of the corridor had the lowest suitability for reintroductions. The rest of the savanna area of the park was similar, being of poor to moderate suitability for reintroduction.

We applied equal weighting to the positive (suitable habitat) and negative (high poaching risk) factors. We have used this approach because there was no objective way to assigning any other weighting. We consider that by presenting the separate analyses of habitat (Fig. 3) and poaching risk (Fig. 5b), managers are in a position to make their own judgment, independent of the results in Fig. 6.

4. Discussion

4.1. Rhinoceros population size

The aerial transects were designed to monitor all ungulate species and only covered 5% of the system. In contrast, total counts were designed to detect large herds of elephant and buffalo and flight lines were spaced too far apart accurately to count solitary animals such as the rhinoceros. Yet, both data sets still provide a reasonable estimate of the black rhinoceros population prior to extirpation in 1977 despite large undercounting errors. Our estimate of 450 rhinoceros is similar to
approximately 477–782 animals obtained by extrapolation from Olduvai to the Serengeti woodland (Frame, 1980). It is suspected that the Frame approximation is high since the narrow, densely vegetated Olduvai Gorge naturally supported a higher rhinoceros density than the more sparsely vegetated Serengeti (A.R.E. Sinclair, pers. obs). Considering both estimates, the target population for future conservation would be approximately 500 animals. Despite the fact that the counting studies discussed herein were not properly designed for high accuracy counting of rhinoceros populations, the present work highlights the value of such long term datasets for conservation.

4.2. Habitat suitability

The habitat suitability models identified a combination of factors that determined rhinoceros distribution, namely denser vegetation relatively close to water. These are the habitats that rhinoceros are known to prefer (Goddard, 1970; Conway and Goodman, 1989), and they occur largely in the western side which is at lower elevation with higher rainfall in drier months such as June. Rhinoceros habitat preference generally explains the important factors in the models. The models were based a number of assumptions such as the interpolation of precipitation data or the distribution of poaching data (see below). We explored the sensitivity of these assumptions by using cross validation and found the results indicate that the predicted values were consistent with the true values.

In the present work, historical observations of the rhinoceros population were correlated with contemporary vegetation data to determine suitable habitats. While caution is warranted when comparing data from different time periods, large vegetation changes (primarily the precipitous decline in mature Acacia woodlands) occurred well before 1970 allowing appropriate data matching. In the Masai Mara National Reserve the vegetation in 1966 was largely open grassland with thickets of Croton and riverine forest (Dublin, 1995) and have not changed appreciably since 1970. Acacia woodlands have regenerated (Packer et al., 2005; Sinclair et al., 2007) but these

Fig. 5b – Contours of poaching risk derived from the combination hunters entering the protected area and animal offtake rates. Maps standardized to range from 0 to 1. For Masai Mara National Reserve only data from areas where hunters entered the protected area (Fig. 4b) were used.
were not used by rhinoceros due to the lack of broad-leaved shrub foliage (Muya and Oguge, 2000; Sinclair, pers. obs.). Thus, the current habitats suitable for rhinoceros existed in their present form in the 1970s.

In the Serengeti National Park, the most suitable habitats for the rhinoceros were the *Croton* thickets of the north and the riverine forests of the north and west. These same habitats have been identified as preferred by rhinoceroses in Ma-sai Mara National Reserve (Dublin, 1995; Walpole et al., 2004). These preferred habitats span the Kenya–Tanzania border. Therefore trans-border cooperation will also be necessary for rhinoceros security (Walpole and Bett, 1999; Walpole et al., 2001; Thirgood et al., 2005).

### 4.3. Poaching risk

Poaching risk was assessed from human population density, locations of people found in the park, and rates of offtake of animal species commonly hunted by people. While this analysis is a good stating point from where to assess poaching risk, our current analysis rests on the assumption that while rhinoceros densities are low, organized poaching gangs are not a security concern. After reintroduction and the potential increase in rhinoceros density, the threat of illegal harvesting of rhinoceros by organized gangs will increase. Our poaching analysis provides an idea of where poaching risk from bush hunters is occurring. It is unlikely that organized gangs will follow similar patterns. Currently we have limited understanding of the patterns and habits of these organized poaching gangs. While this analysis provides guidance for the best potential locations for reintroduction based on the current threat to the animals, after reintroduction these threats will change and these changes should be anticipated. Long term planning will need to include frequent reevaluation of poaching risk and commitment from government agencies, non-government organizations and private stakeholders to provide the funding necessary to protect rhinoceroses.

### 4.4. Areas for reintroduction based on habitat suitability and poaching risk

While there are many factors that need to be considered for a successful reintroduction (Kleiman, 1989; Sarrazin and Barbault, 1996), we focused on only two, habitat suitability and poaching risk. By combining these, the optimum area for rhinoceros reintroduction is the central and eastern part of the

---

Fig. 6 – Areas for reintroduction obtained from merging poaching risk (Fig. 5b) and habitat suitability (Fig. 3). Dark areas predict better areas for reintroduction.
North zone. However, once released rhinoceroses are likely to move towards the most suitable habitat which lies in the high poaching risk area as observed by Tatman et al. (2000). Therefore this provides a conservation concern.

4.5. Conservation and management implications

This study exemplifies how long term systematic monitoring can provide baselines for conservation objectives. In the Serengeti-Mara Ecosystem the original objective in the 1960s was to obtain monitoring data on population sizes and their trends. Managers could not have foreseen the catastrophic collapse of the rhinoceroses and other species a decade later due to political and economic events. The prior data have been invaluable in allowing the analysis of suitable habitat. Contemporary data on poaching has allowed analysis of poaching risk. The combination allows advice on suitable locations for release of rhinoceroses. In summary, long term monitoring data both documents the system and allows the assessment of ecosystem change due to disturbances. These provide the insurance policy for unforeseen events.

Acknowledgements

This work was made possible by the contribution of data from many sources; Michael Coughenour provided precipitation modeling software and was instrumental in the resurrection of the 1969–1972 animal census data, International Livestock Research Institute provided the Kenya human population data, Martin Loibooki provided the Tanzanian human population census data, the Tanzania Wildlife Research Institute and the Frankfurt Zoological Society permitted us to use the current animal census data, and Michael Anderson, Jan Dempewolf, Denne Reed, Suzanne Serneels and the lead author created the vegetation map. We also acknowledge Mike Norton-Griffiths for organizing the 1969–1972 animal censuses. We are grateful to Tanzania National Parks and Tanzania Wildlife Research Institute for their continued support of the Serengeti-Mara Ecosystem the original objective in the 1960s was to obtain monitoring data on population sizes and their trends. Managers could not have foreseen the catastrophic collapse of the rhinoceroses and other species a decade later due to political and economic events. The prior data have been invaluable in allowing the analysis of suitable habitat. Contemporary data on poaching has allowed analysis of poaching risk. The combination allows advice on suitable locations for release of rhinoceroses. In summary, long term monitoring data both documents the system and allows the assessment of ecosystem change due to disturbances. These provide the insurance policy for unforeseen events.

Appendix A. Model crossvalidation

The data was split into $K = 10$ parts consisting of sample count locations. For each part, the models were fitted to the remaining $K - 1 = 9$ parts of the data. The fitted model was used to predict the part of the data removed from the modeling process. This process was repeated 10 times so each sample plot was excluded from the model fitting step and its response predicted. The prediction errors can then be inferred from the predicted minus actual values. Repeating this process over many deleted subsets allows an assessment of the variability of prediction error.

To evaluate the reliability of the model, we computed various measures of prediction error. Prediction bias was calculated for each validated dataset as a percentage of the true value. Accuracy was measured by the mean absolute error, which is a measure of the sum of absolute residuals (i.e. actual minus predicted) and the root means square error, which is the square root of the sum of squared residuals. Small mean absolute error values indicate a model with few errors, while small values of root means square error indicate more accurate predications on a point-by-point basis (Schloeder et al., 2001). To assess the estimation uncertainty in the model (Isaaks and Srivastriva, 1989), we calculated the variance associated with an estimate at a new location, $\text{var}(\hat{Y}_r)$ as

$$\text{var}(\hat{Y}_r) = \text{var}(\hat{h}_r) + \text{var}(\hat{Y}_r) + \text{var}(\hat{c}(T_j))$$

where $\text{var}(\hat{h}_r)$ reflects the uncertainty associated in estimating the parameters of the Poisson regression model, $\text{var}(\hat{Y}_r)$ reflects the random variation at a new location, and $\text{var}(\hat{c}(T_j))$ reflects the uncertainty in estimating the error associated with the regression model from the binary regression tree at terminal node $T_j$. The consistency between the estimation error variance and the observed estimation errors (i.e. true errors), $\epsilon_i = (Y_i - \hat{Y}_i)$, was calculated using the standard means squared error (Hevesi et al., 1992)

$$\text{Standard means squared error} = \frac{1}{n} \sum \epsilon_i^2.$$

The estimation error variances were assumed consistent with the true errors if the standard means squared error fell within the interval $[1 \pm 2(2/n)^{1/2}]$ (Hevesi et al., 1992). Paired $t$-tests ($\alpha = 0.05$) were used to test for differences between the mean estimation errors and zero. The estimation error variances were also used to construct 95% prediction intervals around individual estimates. Coverage rates were calculated as the proportion of individual confidence intervals that contained true values. All statistical analysis was preformed using Splus (Insightful, Inc.).

The overall contribution of the model in describing habitat suitability was 64%. The regression models alone explained 29% of the observed variability in habitat suitability. The binary regression tree accounted for an additional 35%. Prediction bias was nominal. Minimum, maximum, and quartile values showed that estimated and observed value distributions were similar. The mean estimation errors were not significantly different for zero ($p$-value > 0.050). The mean estimation errors was smaller than the root means square error for all models indicating that in general, the model was more accurate in predicting regional or global means than on a point-by-point basis.

Standardized mean square error result showed that the computed estimation error was statistically consistent with the true errors for the models, as it was within the interval [0.72–1.28] (Hevesi et al., 1992). This suggests that estimation error variance could be used to assess estimates of uncertainty for new observations. The 0.95 confidence coverage rate was 0.94. This suggests that prediction intervals constructed using the estimation error variances are sufficiently large enough to insure a 95% prediction interval around our estimates.

Habitat Suitability Model

Number of Observations: 526

$R^2$: General Linear Model – 0.29; Regression Tree – 0.35; Total – 0.64
Variables Used in the Model

**General Linear Model:** Elevation, January, June, November and October precipitation.

**Binary Regression Tree:** Site severity index, elevation, variation in elevation, distance to permanent water, mean annual precipitation, vegetation type, and variation in percent cover of trees.

### Summary statistics of observed and estimated habitat suitability from 10-fold cross-validation

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Observed</th>
<th>Modeled estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>CV%</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>First quartile</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Median</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Third quartile</td>
<td>2.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Maximum</td>
<td>9.0</td>
<td>6.8</td>
</tr>
<tr>
<td>Bias%</td>
<td></td>
<td>6.1</td>
</tr>
</tbody>
</table>

1 CV% - coefficient of variation.

### Summary statistics of estimation errors of the habitat suitability model from 10-fold cross-validation

<table>
<thead>
<tr>
<th>Statistic</th>
<th></th>
<th>Observed</th>
<th>Modeled estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td>0.07</td>
<td>1.1</td>
</tr>
<tr>
<td>IQR</td>
<td></td>
<td>1.40</td>
<td>1.0</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td>1.01</td>
<td>0.0</td>
</tr>
<tr>
<td>RMSE</td>
<td></td>
<td>1.46</td>
<td>0.0</td>
</tr>
<tr>
<td>SMSE</td>
<td></td>
<td>1.00</td>
<td>0.0</td>
</tr>
<tr>
<td>0.95 confidence coverage rate</td>
<td></td>
<td>0.94</td>
<td>0.94</td>
</tr>
</tbody>
</table>

1 IQR = interquartile range, MAE = mean absolute error, RMSE = root mean square error, SMSE = standardized mean square error.

Frequency distribution of the estimation errors of habitat suitability from 10-fold cross-validation.

References


