SCHOOL OF MATHEMETICS AND STATISTICS

Zihan Xu

Statistical analysis for landscape covariate effect on wildlife

MSc Statistics

Student Number: 190008505

Supervisor: Prof. David Borchers

School of Mathematics and Statistics

Mt 5099 MSc Dissertation Project

2020. St. Andrews

Acknowledgement

We are grateful to Jeannie Hayward of the Cape Leopard Trust for providing us with the data used in this dissertation.

Declaration

I hereby certify that this dissertation, which is approximately 8000 words in length, has been composed by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a degree. This project was conducted by me at the University of St Andrews from July 2020 to August 2020 towards fulfilment of the requirements of the University of St Andrews for the degree of MSc Statistics under the supervision of Prof. David Borchers.

Signed: ____________________ZihanXu_________

Date: ___________08/22/2020_________________

Content

Abstract

Landscape components could provide important information for ecologists to locate hotspot habitats in habitat protection. Based on spatial capture-recapture theory, the encounter model is developed and applied in this study to analyze the roles of each landscape component. By using Euclidean distance between pixel locations and individual activity centroids, this model calculates the count of individual for each pixel, with extension of landscape covariates added to the model. GLM Poisson encounter model is adopted to estimate the association between two landcover components, with individual space usage explored for nine leopards near South Africa. The results of this study's model can demonstrate how elevation and landcover components can affect leopards' space usage. Because this model is based on the assumption of independent individual space usage, it will not work if individuals' movements in the habitats go against the assumption. Thus, further improvement in the future could be made by analyzing individual movements. An improved model could provide additional information based on individual movements across pixels.

Key Words: Encounter Model, Landscape covariate, SCR model.

Introduction

Animal habitat protection is a very important topic, because habitat destruction could cause bad effect in humans (Primack 2006). In order to allocate human facilities for financial benefits while protecting habitats for wild animals, it is necessary to deliver moderate development in wild areas. A good way for this practice is to impose habitat conservation during human development near wildlife habitats. In this respect, the first matter needed to be considered is how to identify priority of habitat areas' conservation and protection. But it is pretty hard and complex to determine the size, type and location of habitats in biology field (Hierfl 2008). Traditionally, hotspot areas are identified by observations, which are, however, difficult to precisely estimate the locations of habitats, because many animals, such as leopards and condors, tend to move in so large spaces that some locations in their activity area are not very important to be considered. In order to pinpoint the hotspot regions for wild animals efficiently, some landscape parameters related to habitats could be estimated and analyzed to help search the really important habitats for animals. For instance, pandas tend to live in the areas where bamboo exists; thus, it is reasonable to estimate panda habitats according to the distribution of bamboo. Such landscape parameters highly related to individual's movement and density could be identified and referred to during habitat protection. To this end, the relationship between different environment elements and animals should be estimated.

Applied widely in animal population density study, the spatial capture-recapture model uses the encounter probability model to estimate the encounter probability via a function of Euclidean distance and make spatial analysis in density research (Borchers & Efford *et al*. 2009; Gardner *et al*. 2010; Gopalaswamy *et al*. 2012; Royle *et al*. 2013a, b). Spatial capture-recapture (SCR) model could avoid the bias in density study resulting from undefined effective trapping areas; and it could estimate animal density based on repeated animal observation data from multiple traps, a cost-efficient approach compared with individual observation data during a long time period. By introducing a new model related to spatial capture-recapture model, Royle (2013b) solves the problem that the encounter probability model under SCR model may make the home range of animal activities too stationary and too symmetric due to lack of landscape effect analyses; further, by integrating landscape covariate parameters into the function of encounter probability model, Royle's SCR model could both analyze animals' space usage via environment influence and predict animals' population density via landscape connectivity.

In this study, the encounter model in SCR model is used for a different purpose: to find out the relationships between animal space usage and landscape covariates in SCR model. The aim in this paper is analyze how do landscape components affect individuals in this area. By fitting encounter model based on Royle's SCR model (2013b), it is possible to identify the effects of some landscape components to animal individuals space usage. Some important information about landscape effects could be acquired from the estimated coefficient of each landscape parameter in the encounter model. In summary, landscape effect analysis under the method developed in this study could solve three important issues: (i) it can determine what kinds of environment components are important for animal individuals to live in certain habitats and what kinds of landscape elements are not important; (ii) it can estimate how these kinds of landscape components influence animals' activity selection (in positive effects or negative effects); and (iii) it can calculate the extent of the impact on wildlife by these landscape covariates (in strong effects or low effects). The key methodology in this study is to adopt the spatial capture-recapture model to explore how environment components would influence animals and how animals leverage their habitats.

The following sections will introduce the SCR model for density study and discusses the basic SCR model and Royle's (2013b) SCR model with landscape covariates in the encounter probability model. Then the method to analyze landscape parameters based on encounter model is described in "Encounter model" section, followed by illustrating the application of this new model in exploration of the landscape effect on leopards in their African habitats. Finally, GPS location data of nine different leopard individuals' activities are recorded by GPS collar on each leopard, with some spatial-referenced data covering that area (for elevation and landcover). We use our method to find how each landscape elements have affected the nine leopard individuals' usage of their habitats, with some suggestions offered for protecting these leopards' habitat.

Literature review

After review of literatures, how the theories and insights have been woven is revealed, including: the efficient methods in animal population density study, the problems and solutions in classical animal density study and spatial capture-recapture study, the encounter probability study, and the effective methods to connect animal density researches with landscape covariates. This chapter will review the existing studies relevant to all threads of animal population density research. Each of the threads is related to the topic of this article and has some space to be improved. Royle (2003) proposes a new way in animal density research to deal with heterogeneity problems and achieve financial benefits. Efford (2004), Borchers (2008) and Royle (2008) have offered new methods to solve the problem of undefined effective trapping areas. Rolye (2013a, b) analyzes the animal space usage in spatial capture-recapture model as well as environment effects on animals. This review shows that the traditional capture-recapture model in animal density study fails to consider the landscape effect in density study, an issue this research is aimed to solve.

Animal population density is one of the most important parameters in animal population study (Krebs 1985; Turchin 1998). And collecting animal population data from traps might be the most common method in animal population density analysis (Thompson *et al*. 1998). In addition, the capture-recapture method is normally used in population size study (Borchers & Efford 2008). Royle (2003) introduces a method to estimate the occupancy rate when heterogeneities in trapping probability exist. The basic idea of the method is that the detection probability could be affected by the variance of abundance. This method can facilitate the estimation of animal abundance distribution by making models of heterogeneity detection probability, and it can accept animal data from repeated trapped records instead of unique individual population documents (Royle & Nichols 2003). By using this method, an animal population research could be made several times, thus saving a large amount of money during recording observations. In the capture-recapture method, by dividing effective trapping areas, population sizes could be converted to or treated as density; however, effective trapping areas may be very hard to calculate (Efford 2004). Dice (1938) suggests the distance from the traps could be used to calculate the effective trapping area. There also exist some other ways to get trapping areas, but none of them are commonly used or strongly reliable (Efford 2004). Several other methods to avoid effective areas in population study, like trapping webs (Anderson et al. 1983), require specific trapping conditions and have weak reliability for population density in most cases (Efford 2004). Parmenter's study (2003) about the density estimation of rodents from 11 data sets compares the grid-based density estimation methods and the web-based density methods by using given density value datasets. Both methods in his study have result that meet the assumption of each method. This study finds that most web-based methods show better accuracy than most grid-based methods; the best web-based methods deliver comparable performance to the best grid-based methods; the method using "effective trapping area" delivers poor performance in rodent density studies, but the application of rodent movement distance can significantly improve the performance of density estimation. The web-based methods seem with better performance than the grid-based one, but the former requires several assumptions and well-defined effective trapping areas. Therefore, a good performance of trapping web methods entails good measured trapping areas, which are actually hard to estimate in most studies, however.

To solve the problem of no efficient ways to define and estimate effective sampling areas, Efford (2004) proposes a new method to estimate density from capture and recapture animal population data without calculating effective trapping areas. Specifically, it uses a two-parameter spatial capture function to make a simulating model and do simulation and inverse prediction to estimate population density. In experiments, this method provides an unbiased estimate of population based on simulation data (Efford 2004). However, some problems appear in this method when trap saturation level becomes high enough and when trap competition exists; in addition, this method has some limitations in model selection and covariate analysis (Borchers & Efford 2008).

As an approach very similar to Efford's, the Spatially Explicit Maximum Likelihood Methods developed by Borchers (2008) solve the problem that capture-recapture method is unable to analyze the spatial natures of population density. Under the capture-recapture method, spatial components could affect the trapping progress, because traps may, in a high probability, record the animals close to the traps than the animals far from the traps. This problem has not been solved in the previous population estimate studies, while it may deliver results not reliable enough for spatial density analysis (Borchers & Efford 2008). The new method could achieve spatial nature analysis by estimating animal spatial probability from animal locations to trap locations based on the maximum likelihood. With spatial components added into the model, the density estimation could be "well defined," because the population abundance could be defined.

Another problem with animal density study for close populations is the existence of population movement, mostly in a short term, which will cause effective population areas to change and make the estimate result not strongly reliable (Royle & Young 2008). Based on capture-recapture data, Royle proposes a new method (2008), which involves a hierarchical model to calculate sample spatial unit areas, and then individual activity centers and individual activity movements could be estimated by the hierarchical model. The density of individuals could be estimated by this method, and thus the problem of population movement could be solved.

Among these three methods, one common component is the application of the spatial capture-recapture method, the elementary theory of which is to use a function containing the distance between individual animals' location center and the trap location to acquire the encounter probability (Borchers & Efford 2008). As such, it is possible to finish spatial population density research. Although spatial capture-recapture models are a new class of methods, their use is relatively widespread in exploration of population density and movement estimation (Borchers & Efford *et al*. 2009; Gardner *et al*. 2010; Gopalaswamy *et al*. 2012; Royle *et al*. 2013a, b). While using a new spatial capture-recapture model to find animal density of close-population with stable home range, Borchers and Efford (2009) also introduces a new longtime single sampling interval, which has some financial benefits as compared with previous sampling time intervals. Then, Gardner's hierarchical spatial capture-recapture model (2010) uses spatial explicit components of black bears detected by traps to track them. Gopalaswamy (2012) further makes a comparison capture-recapture research on tiger density, tiger photographic data analysis and tiger fecal sample analysis, finding that making inference on parameters from more information could improve the model. SCR model could be enhanced by adding more landscape information to the model. The widely application of spatial capture-recapture (SCR) model comes from its ability to solve certain critical problems in non-spatial capture-recapture methods (Royle *et al*. 2013a, b), such as the heterogeneity in encounter probability from juxtaposition of individuals with traps and ill-defined sampling areas (Borchers 2012; Royle *et al*. 2013a, b). The auxiliary spatial information produced in capture-recapture research is also a benefit of using the spatial capture-recapture method.

Royle (2013b) points out that, despite the wide application of spatial capture-recapture method in population density analysis, the encounter probability model makes the estimation of space too stationary and too symmetric to represent animals' spatial resource selection. The encounter probability model, which most capture-recapture models are based on, uses Euclidean distance to calculate the encounter probability by simple functions in spatial capture-recapture model; and some different encounter models have been proposed over resent year, including binomial, Poisson and multinomial encounter models (Royle *et al*. 2013a). The encounter probability model implies a stationary and symmetric animal home range, but this assumption is unreasonable and ungrounded, because influenced by environments, animal often make some improvised movement decisions. For example, water buffalos tend to move to areas with a plenty of water, while leopards prefer those areas which are their prey habitats. Influenced by resources, the function in the encounter model can't analyze the landscape effects or get the encounter probability. Although such stationary distribution could show good performance with some sparse data, these simple encounter probability models can't represent animal home range size and shape in most cases, due to animals' uneven spatial resource selection (Royle *et al*. 2013b). The location where an animal lives and the state of its surrounding environment could affect the animal's spatial usage, so the encounter model should contain irregular and non-stationary properties on animal home range and could deal with local resource effects.

Landscape connectivity, which is defined as the level of landscape components affecting animal movements (Tischendorf and Fahrig 2000), is recognized widely as an importance element in the population viability (With and Crist 1995). In ecological research, the landscape connectivity in animal density effect is one of the most important properties (Royle *et al*. 2013a). The existing researches typically estimate landscape connectivity values based on the researcher's subjective opinions or some temporary methods (Adriaensen *et al.* 2003; Beier *et al.* 2008; Zeller *et al.* 2012). Without proper methods to predict animal density and landscape connectivity values simultaneously, it becomes very hard to predict the effect of landscapes on animal density. This problem could cause huge bias in population density estimation.

Royle (2013a) has developed a model based on spatial encounter probability with alternative distance, which could represent connectivity. Called by Royle the "ecological distance", this distance is related to the least-cost path with a cost-weighted distance metric, and could define the distance between an animal activity center and traps in cost-weight metric. The ecological distance could be calculated based on the cost of moving between different cells in sample data, while considering some environment variables (e.g., elevation, snow cover, landcover and range limitation) that may affect animal movement between two cells during the calculation (Cushman *et al.* 2006; Schwartz et al. 2009; McRae and Beier 2007). By using such a distance, the landscape structure's influence could be added to the spatial density analysis; and the parameters of cost-weighted distance could be calculated and used directly in spatial capture-recapture method to connect the animal spatial density with landcover covariates. According to Royle's simulation study (2013), the model using Euclidean distance would produce huge bias, while the model with least-cost path distance deliver good results. However, one problem of this method is that it cannot estimate the precise location of animals' movement paths and decisions when they move in different landscapes (Royle *et al*. 2013a).

The second model that could deal with the spatial capture-recapture process with animals' space usage is proposed by Royle (2013b), who improves the spatial capture-recapture model by adding one or more clear landscape covariates to the encounter probability model, thus extending the model with space usage and resource selection. In this way, the elements that have not been presented in the existing encounter probability models, such as animal home range size and shape, could be included in Royle's new model, which can then relate capture-recapture analyses with the trend of animal space usage and can also avoid bias during estimating process, especially when landcover covariates may influence animals' spatial usage and selection (Royle *et al*. 2013b).

According to previous researches on animal population density (Efford 2004; Borchers & Efford 2008; Royle & Young 2008; Royle *et al*. 2013a, b), the spatial capture-recapture method could be used in animal population density analysis, because the method has financial benefits, compared to previous trapping methods, and could solve the problem in classical non-spatial methods. In addition, the application of encounter probability model in spatial capture-recapture model could address two big problems in traditional population density analysis, i.e., to avoid the problem of undefined sample trapping area. and to deal with spatial elements in population density study. However, the encounter probability model might cause some problems in estimation from animal density data with landscape effect, since the encounter model could not connect the animal home range with the surrounding landcover situations. Because of the limitation in analyzing animals' surrounding environment components in the encounter probability model, the animal home range estimations by the model's function would highlight the stationary and symmetric nature and make the estimation unable to represent environment effects in density study. Thus, it is necessary to add landscape covariates in spatial capture-recapture model, so as to reveal the connection between animals' spatial components and the landcover resource elements. This approach could improve the estimation of density when landscape covariates have influences on animals' spatial movement and usage, while providing the landscape effect analysis in animal population study.

Spatial capture-recapture method

The spatial capture-recapture method is a new method to estimate animal population density from individual capture-recapture location data (Royle *et al*. 2013b). Its key concept is to use a function based on the distance between trap locations and animal activity centers to identify the encounter probability. The trap location (x) is each individual trap's location in capture-recapture data, and the activity center (s) could be the home range of an individual, which is calculated by taking the average of that individual's trapping locations during a specified time period (Royle *et al*. 2013a). The introduction of spatial components in the capture-recapture method would avoid bias in estimating abundance by modeling the least trapping individuals, and facilitate the model to deal with the conditions when the capture probability depends on locations. Models based on spatial capture-recapture methods could solve all problems in close-population capture-recapture studies, such as the problems with observable or unobservable heterogeneity and behavior response (Borchers & Efford 2008).

Many models are available to estimate the encounter probability in spatial capture-recapture study, including Poisson model, binomial model and multinomial model (Borchers & Efford 2008; Borchers & Efford *et al*. 2009; Royle *et al*. 2009). In binomial encounter model, we suppose there has j traps and thus x_j trap locations; and we also suppose i observation individuals in the capture-recapture data; thus, we have a total of N_{ij} pieces of observation records based on number of individuals and traps. For individual i, the vector $N_i = (n_{i1}, n_{i2}, \ldots, n_{ij})$ forms its encounter history. From the observation n_{ij}, we could make a gaussian model that contains a binomial function to get the encounter probability p_{ij}, and this model is a standard encounter probability model (Borchers & Efford 2008).

$$log(p_{ij}) = \alpha_0 - \alpha_1 \cdot d_{ij}^2. \qquad \text{(eqn 1)}$$

This function could also be written as (Royle *et al.* 2013b):

$$p_{ij} = \lambda_0 \cdot exp(\frac{-d_{ij}^2}{2 \cdot \sigma^2}) \qquad \text{(eqn 2)}$$

where d_{ij} is the Euclidean distance between trap location x_j and individual activity center s_i,

$$d_{ij} = ||s_i - x_j||$$

and then let α_0=log(λ_0) and α_1= 1/(2×σ^2), this will be used in Encounter model later.

This model is based on a basic assumption: the Euclidean distance in the standard encounter model is stable and can't be influenced by other variables, such as environment components. In other words, it assumes that all individuals here have a stational and symmetric home range and individuals move in same circular spatial format in their habitats, but this assumption is irrational and unreasonable under real conditions. In fact, landscape components and habitat conditions will affect animal activities and cause them to change their space usage. As such, an irregular and non-stational home range is required in population density study, so as to accommodate real habitat environments. The new model should be able to react sensitively to different spatial landscape conditions and estimate individual space usage based on the landscape covariates from surrounding areas.

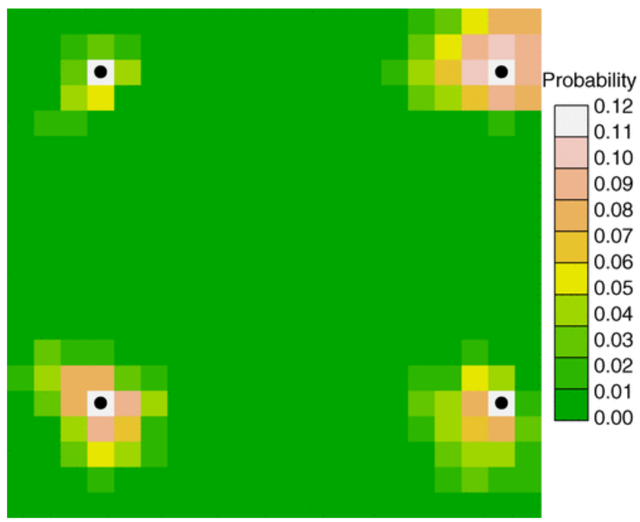

Figure 1

Typical home ranges for four individuals in a fragmented landscape (Royle *et al*. 2013a).

This plot shows 4 individual activity centers and their encounter probability in pixels. From the color of each pixel, we could read the home range of each individual. The activity centers sometime are not reflected in the pixels to indicate that individuals use the most intensive or the center of their activity areas. The plot of one individual trap observation shows its activity center and the frequency of usage for each pixel.

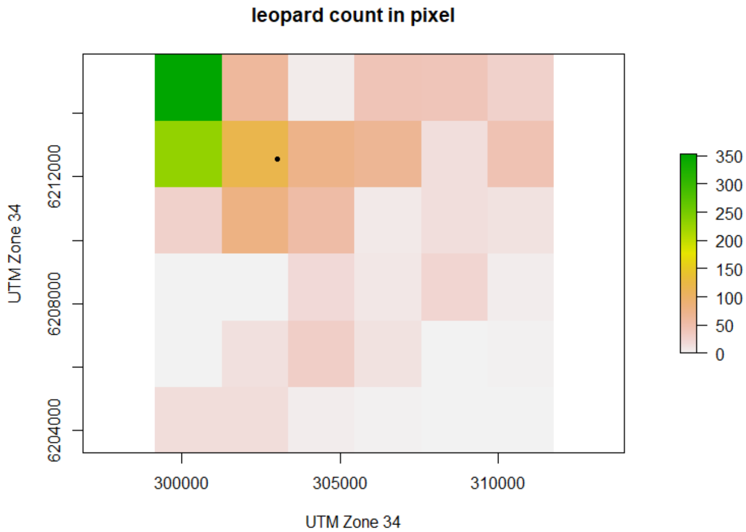

Figure 2

One leopard home range plot with its activity center.

Model of space usage

The new method focuses only on spatial usage of individuals instead of their activity patterns, so this study only needs to model the results of observation data from traps and does not need to analyze animals' movement process (Royle *et al*. 2013b). Here, we define data as raster format for individual locations and landscape covariates. Suppose there are n pixels in the raster data, and the set of $x_j = (x_1, \ldots, x_n)$ is the pixel coordinate in the data and could be recognized as the trap location which described in Section "Spatial capture-recapture method". Individuals' space usage in their habitats is defines as their location record in pixels from the observed raster data. Then this study defines z as the covariate of landscape (elevation or landcover) in this raster. In addition, this study could add more than one landscape covariate, as discussed later.

In multinomial assumption, let $p(x_j)$ represent the probability of individual to use j pixel:

$$P(x_j) = \frac{exp(\alpha_2 \cdot z(x_j))}{\sum_1^n exp(\alpha_2 \cdot z(x_j))} \quad \text{(eqn 3)}$$

which is a standard RSF model, according to Manly (2002). Coefficient α_2 represents the influence of landscape covariate on the probability of pixel usage, which will increase as $z(x_j)$ increase when the positive landscape coefficient α_2 exists. Then the equation 3 is extended to make a probability model using each pixel location and the activity center of individuals i (Johnson *et al*. 2008; Forester *et al*. 2009):

$$P(x_j|s_i) = \frac{exp(-\alpha_1 d_{ij}^2 + \alpha_2 z(x_j))}{\sum_1^n exp(-\alpha_1 d_{ij}^2 + \alpha_2 z(x_j))} \quad \text{(eqn 4)}$$

where x_j is each pixel coordinate, s_i is the activity center of individuals, and d_{ij} is defined as the Euclidean distance from the activity center to pixel x_j. This function describes the relationship between the encounter probability for individuals (also the probability of pixel usage with an activity center) and the distance of pixel to the activity center, with additional landscape covariate. When α_1 is positive, as the distance from pixel to activity center increases, the encounter probability will decrease; and when positive α_2 exists, as the value of landscape covariate $z(x_j)$ increases, the probability will increase. Thus, the equation could be simplified to:

$$P(x_j|s_i) \propto exp(-\alpha_1 d_{ij}^2 + \alpha_2 z(x_j)) \quad \text{(eqn 5)}$$

If α_2 is set 0, it represents there has no environment effect on the encounter probability, so equation 5 will change to:

$$P(x_j|s_i) \propto exp(-\alpha_1 d_{ij}^2).$$

Encounter Model

There has another model relate to SCR study, not to estimate the individual encounter probability but estimate the observation count for each pixel (x_j), j represents number of pixels. The hierarchical model based on assumption that x_j (j pixel) with i individual has a Poisson distribution with mean value (λ_{ij}), where the expected value of Nij (pixel j and individual i) is $E_{ij} = \lambda_{ij}$. The individual activity center is s_i, and the Euclidean distance $d_{ij} = ||x_j - s_i||$. The model for expected count is (Chandler & Royle 2013):

$$N_{ij} \sim Poisson(\lambda_{ij}),$$

$$E_{ij} = \lambda_{ij} = \lambda_0 exp(\frac{-d_{ij}^2}{2\sigma^2}). \quad \text{(eqn 6)}$$

This formula could be transferred to:

$$\lambda_{ij} = exp(log(\lambda_0) - \frac{1}{2\sigma^2}(d_{ij}^2)),$$

where log (λ_0) = a0 and α_1= 1/(2×σ^2) according to spatial capture-recapture model (Royle 2013b). Then we have:

$$\lambda_{ij} = exp(\alpha_0 - \alpha_1(d_{ij}^2)) \qquad \text{(eqn 7)}$$

Then according to Royle's research in 2013 (Royle 2013b), we could add landscape covariate z that affect the expected count λ_{ij} to the model:

$$E_{ij}|d_{ij} = \lambda_{ij} = exp(\alpha_0 - \alpha_1 d_{ij}^2 + \alpha_2 z_{ij}). \qquad \text{(eqn 8)}$$

In Equation 8, α_0 is the intercept, α_1 is the parameter for the distance from j pixel to I individual activity center, and α_2 is the parameter for landscape covariate. This function shows how the distance d_{ij} and landscape covariate affect the individual expect count. When α_1 is negative and α_2 is positive, as the distance and landscape covariate increase, the expect count will also increase.

This study will use this GLM Poisson model based on this equation to model the count from eight leopards' individual location data. After fitting the model for the count, this study could find the relationship between landscape covariate parameters and leopards' space usage. If the coefficient of the landscape α_2 is found negative, that means this landscape component has a negative effect on individuals. The bigger the α_2 value, the stronger the influence of this component on individuals. After analyzing the coefficient, this study could get the information on the connection of animals with their surrounding habitat conditions, so as to help protect their habitats more efficiently.

Application: South Africa leopard study

The GLM Poisson spatial capture-recapture model is applied in this study about nine leopards around their habitats in southern Africa. Researchers name these leopard individuals as BF8, BM12, BM17, BM18, BM21, BM22, BM4, BM5 and BM7. For total 14,528 GPS locations have been recorded by two types of GPS collars that could report GPS information on leopard individuals; and these GPS records begin from September 2012 to September 2014. Expect the x-coordinate and y-coordinate in GPS data, elevation and landcover were also recorded. Therefore, this study has the location for each individual observation and two landscape covariates around individual habitat. The main focus in this study is to explore how these nine leopards use their habitats; and efficient method is applied to analyze landscape covariates (elevation and landcover) in the fitted model: (i) to identify which environment covariate is significant in the leopards' individual space usage, and (ii) to find out the relationship between the covariate and the leopard counts in the record data as well as its extend. We could get the information on landscape covariate from the coefficients and the p-value after fitting the GLM model.

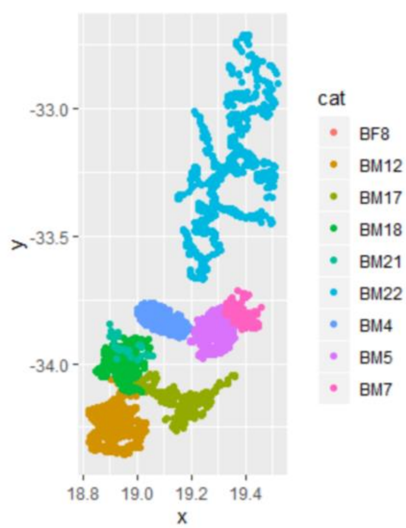

Figure 3

GPS locations of 9 leopard individuals.

To use GLM Poisson SCR model, this study first needs to parse the location data into pixels, the x_j we talked previously in encounter model. After cutting the data into 87x30 areas, a total of 2,610 pixels that contains all GPS location points in the data are acquired, and then this study calculates the total amount of individual observations in each pixel. 0-448 counts for leopards are identified in all pixels. Then in data processing, the missing values in GPS location data are dropped, and the raw GPS information is transferred to UTM Zone 34 format (for zones of leopard's habitats). Then, by taking the average of each of their GPS coordinates, the activity centers of the nine leopards can be identified. These nine activity centers are s_i in SCR model, and i in s_i is from 1 to 9. The pixels cut from the GPS data are plotted with colors to represent the count in each pixel, and nine activity centers are added to the plot.

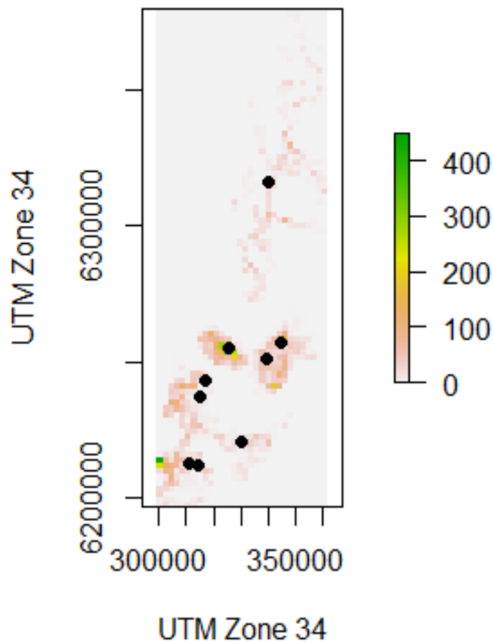

Figure 4

Activity center of nine leopards and the count of leopard for each pixel.

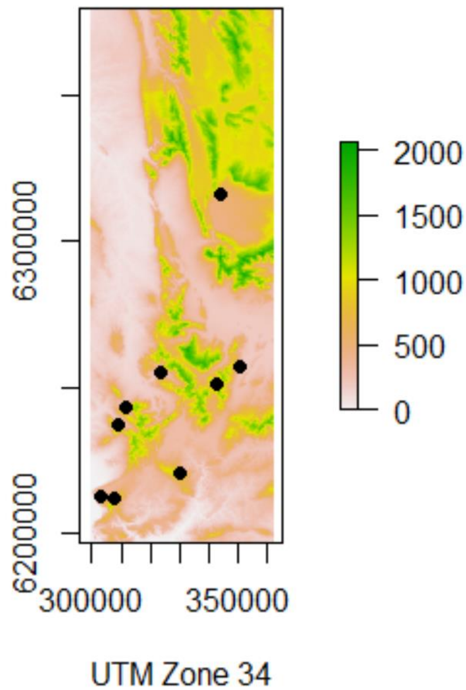

Figure 5

Activity center of leopard and elevation in each pixel.

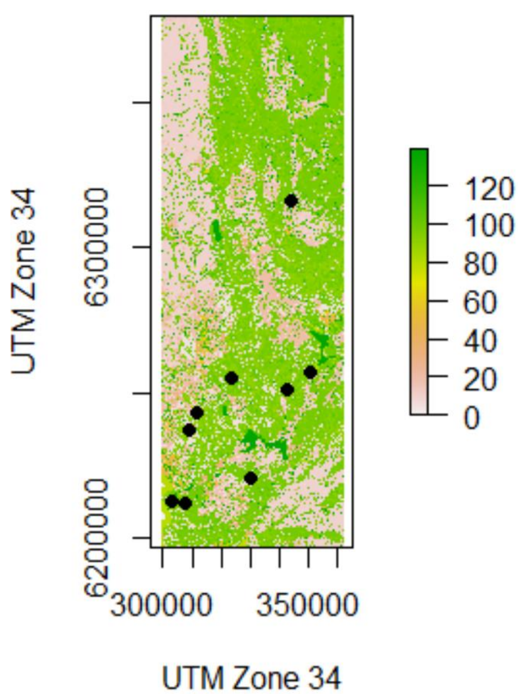

Figure 6

Activity center of leopard and landcover in pixel.

As shown in Figure 4, the distribution of nine activity centers is not symmetric or regular; rather, most activity centers of the leopard individuals are located in the bottom of the area. And the home ranges, as indicated by the color areas around activity centers, show that the nine leopard individuals might have different space usages for different sizes of their home ranges. The eight individuals in the bottom of the area have a smaller home range than the individual on the top of that area. This demonstrates that this study might need to add distance analysis for each leopard individual. Figure 5 shows that most of the leopards' favorite area is at an elevation of less 1,000m, and figure 6 shows that leopards prefer to stay in high landcover regions. The relationship of count and landcover and elevation shows more clearly in Figure 7 and Figure 13. The Euclidean distance from the centroid of each leopard to each pixel location, i.e. d_{ij} in SCR model, is then added to the pixel data frame. Because these nine individuals have records in coincident pixels, it is hard to transfer the location data of all the nine leopards together. Therefore, this study transfers the observations of each of the nine leopards to 2,610-pixel data frames (Each row records the count of leopards in that pixel) separately and then combines these nine data frames together. The first 5 lines of data are shown in appendix.

Basic Model

This study have processed data of 23,490 rows, which connect to 2,610 pixels and nine individuals; each row contains x and y coordinates of pixel locations, the count of leopards in that pixel(x), the individuals' names in that pixel, the distance (d) which is the Euclidean distance of each pixel to individual centroid, landscape covariate elevation (z_1) and landcover (z_2). The basic model used in this study is GLM Poisson model, based on the Poisson encounter model (eqn 8), which contains the distance d_{ij} and two landscape covariates z_1 and z_2, the expected value for each pixel is:

$$E(x|d) = exp(\alpha_0 + \alpha_1 d^2(x) + \alpha_2 z_1(x) + \alpha_3 z_2(x)).$$

Then, we make several models related to the basic model, including models with or without landscape covariates and models with single covariate. Here is the full model list:

Model 1: Basic Model - Basic encounter model with all terms.

Model 2: DModel - Model with only distance parameter.

Model 3: EModel - Model with distance and elevation parameter z_1 (x).

Model 4: LModel – Model with distance and landcover parameter z_2 (x).

During fitting the Poisson model, we find large dispersion values for all these four models; therefore, we change the GLM model family to quasi-Poisson for fitting. For quasi-Poisson, traditional AIC test cannot work in these models; so, we use QAIC to test them in this study. The table below shows the information on the first four models:

Table 1. Summary of fitting models. Parameters include intercept α_0, coefficient of distance α_1, elevation α_2 and landcover α_3. Their QAIC, p-value (p-v) and standard error (SE) are also shown in the table. The data contain 9 individuals and 2,610 pixels for each individual.

Model	α_0	p-v	α_1	p-v	α_2	p-v	α_3	p-v	QAIC
Basic	3.85	<2e-16	-2.29e-4	<2e-16	1.13e-4	0.635	3.19e-3	0.151	466.5
SE	2.59e-1		1.5e-5		2.37e-4		2.23e-3		
DModel	4.176	<2e-16	-2.31e-4	<2e-16					465.2
SE	1.62e-1		1.64e-5						
EModel	4.064	<2e-16	-2.3e-4	<2e-16	1.86e-4	0.436			466.6
SE	0.211		1.54e-5		2.38e-4				
LModel	3.902	<2e-16	-2.29e-4	<2e-16			3.39e-3	0.132	464.7
SE	2.42e-1		1.55e-5				12.25e-3		

The QAIC value shows that Lmodel with landcover covariate and the DModel is preferred. Parameter α_3 shows the individual count has delivered a positive response to landcover. Parameter α_1 shows the distance has a negative effect on leopards' space usage, which is proved by the second plot below. The positive response of landcover parameter seems not be shown by the plot.

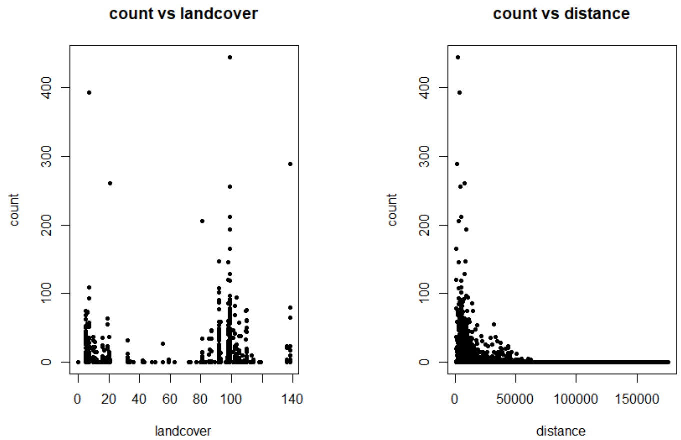

Figure 7

Plots of counts with landcover and distance. The first plot shows that landcover has positive effect on count, and the second plot shows that distance has negative effect on count.

Two problems exist in the table. First, the basic model which contains all landscape covariate parameters is not the preferred model in this table. This means that elevation might have no relationship with the leopards' expected count, which may only depend on the distance and landcover parameters. Second, in the most preferred model of LModel, landcover coefficient α_3 has a p-value that demonstrates this parameter is not significant for the model. The reason for these problems lies in the plots of count and distance. By plotting count and distance for each individual leopard separately (figure 15), this study notices the home ranges of individuals are different; plots for distance and count of individuals of BM22 and BM4 showcase this problem clearly, with BM22 having an activity range of 60,000 and BM4 having a range of less than 10,000. This could cause errors for distance parameters during fitting the model. Therefore, the interaction term of distance with individual factor should be added into this model.

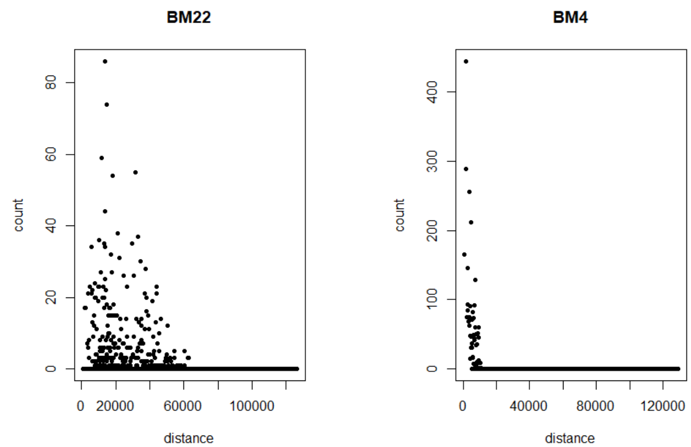

Figure 8

Plot of distance and count for BM22 and BM4.

Interaction Model

An interaction term of leopard individual name factor and the distance is added to the basic model, to as to solve the problem causes by individual home range variation. The new model is:

$$E(x|d) = exp(\alpha_0 + \alpha_1(d^2(x) : cat(x)) + \alpha_2 z_1(x) + \alpha_3 z_2(x)),$$

where $d^2(x) : cat(x)$ is the interaction term and cat(x) is the name of leopard individual here. Because the dispersion parameter is still large, quasi-Poisson model is used to replace Poisson model. Based on the result of dredge function from QAIC test, the model containing the interaction and two covariates is most preferred with the lowest QAIC value. So, this study will use this model as the final model. The estimation of coefficient, p-value and standard error of this fitting model are shown below.

Table 2 Summary of parameters in the basic interaction model. SE is standard error, Estimate is estimate coefficient, and t-value and p-value are shown in the table.

Parameter	Estimate	SE	t-value	p-value
Intercept	5.606	8.215e-2	68.240	<2e-16
Distance	-4.942e-4	1.690e-5	-29.245	<2e-16
BM12	-1.764	1.096e-1	-16.096	<2e-16
BM17	-1.967e-1	1.234e-1	-15.964	<2e-16

BM18	-5.995e-1	1.015e-1	-5.909	3.48e-9
BM21	-3.148	3.526e-1	-8.926	<2e-16
BM22	-2.968	9.906e-2	-29.966	<2e-16
BM4	6.016e-1	9.002e-2	6.684	2.38e-11
BM5	-2.342e-1	9.756e-2	-2.401	0.0164
BM7	-1.409	1.622e-1	-8.688	<2e-16
Landcover	1.688e-3	3.685e-4	4.579	4.70e-6
Elevation	-2.812e-4	4.522e-5	-6.218	5.13e-10
Distance:BM12	2.897e-4	1.865e-5	15.534	<2e-16
Distance:BM17	2.494e-4	1.959e-5	12.732	<2e-16
Distance:BM18	1.686e-4	1.937e-5	12.732	<2e-16
Distance:BM21	3.042e-5	6.734e-5	0.452	0.6514
Distance:BM22	4.149e-4	1.707e-5	24.315	<2e-16
Distance:BM4	3.566e-5	1.943e-5	1.836	0.0664
Distance:BM5	1.798e-4	1.830e-5	9.825	<2e-16
Distance:BM7	1.441e-5	3.337e-5	0.432	0.6659

In this model, all the distances and two landscape covariates show strong significance, indicating that all distances and two landscape components are important for leopards' space usage. The SE (standard error) of these three parameters is reduced in this model. Distance has negative effect on individual space usage, while landcover has a positive relationship with individual counts, as in the previous model without interaction term. It is interesting that, compared to the previous model, the effect of elevation on individual usage is changed to a negative one. So the influence of the interaction term seems to have changed the effect of the elevation.

In model diagnostic, the residual and fitness plot shows a straight line and points tend to have no non-linear trend, with most points are distributed equally across the red line, meaning that there's no discernible non-linear trend to residuals in this model. The nomal-QQ plot has no meaning here, since this model is Poisson distribution. The Residuals vs Leverage plot shows most points stay inside the Cook's distance lines. Four points lie outside the red line, showing high residual and high leverage. The Pearson residuals plot could do nonlinearity test for each parameter in this model. The quadratic smooth lines in residual plot show there is no nonlinearity, since the curvature tends to be zero in all model covariates and the test statistic for coefficients related to quadratic term is not very large.

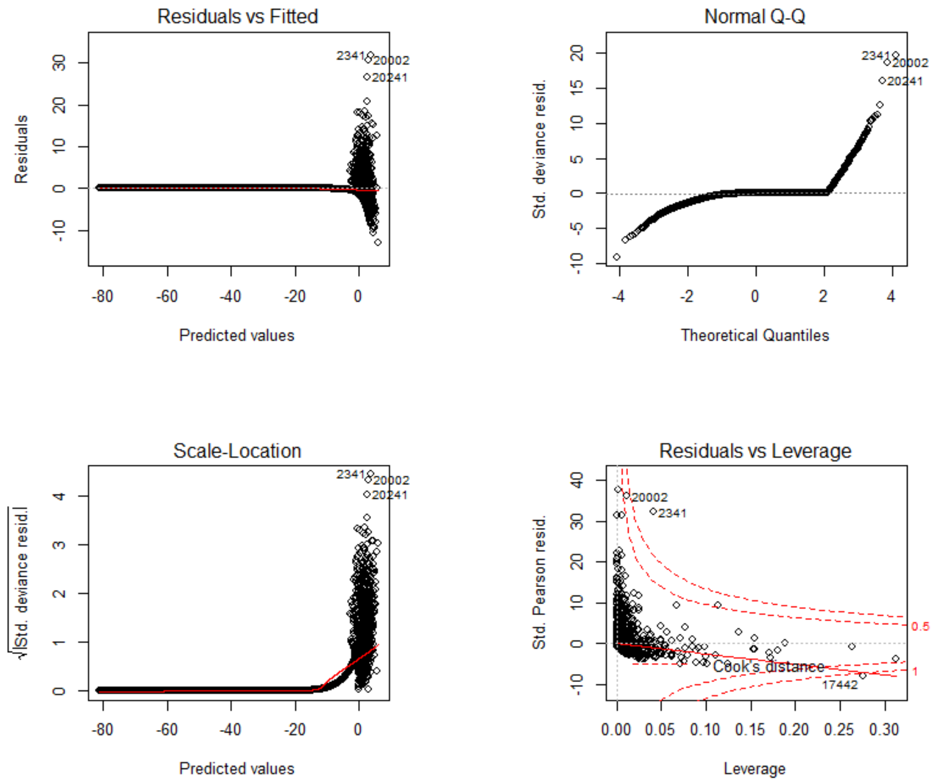

Figure 9

Four diagnostic plots for final model with interaction term

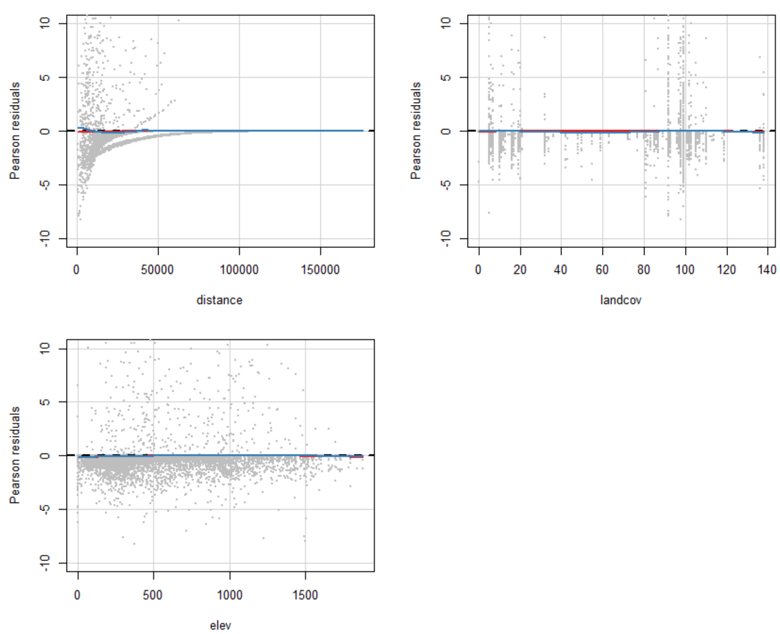

Figure 10

Residual plot for final model parameters without fitted value

Prediction

Predictions of leopard counts are made based on the original data and the final model, including 9 leopards and 23,490 prediction counts in each pixel. If the error between the prediction and the real count is less than 0.5 in each pixel, the result is regarded as accurate, and if larger than 0.5, the result is inaccurate. Tests show that the accuracy of the prediction by this model is around 0.9063. The comparison of data and prediction graphs that describe the count vs each covariate shows the effect of covariates on the count is the same as in the previous model analysis. And the trend in each plot proves the effect on individual count for each covariate, as acquired from the fitting model above. The prediction plot shows that most individuals tend to move in areas where the elevation ranges from around 50m to 700m and landcover has a value from 80 to 120.

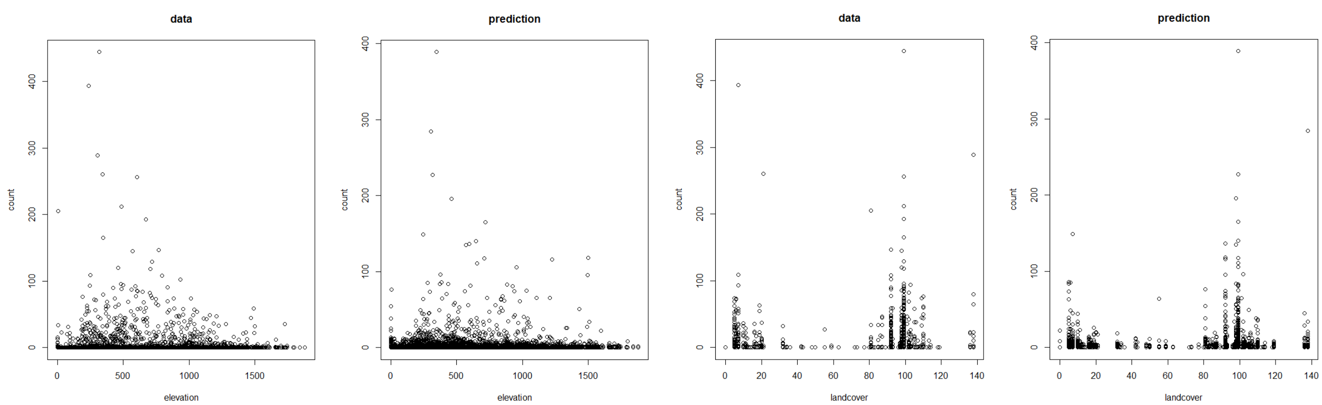

Figure 11

Comparison plot for count vs covariate.

The heatmap based on the predictions could show the effect of model, elevation and landcover component. Below are five pixel plots: the count in pixel plot of original data, the count in pixel plot of prediction, the mean Pearson residuals and pixel plot of elevation and landcover.

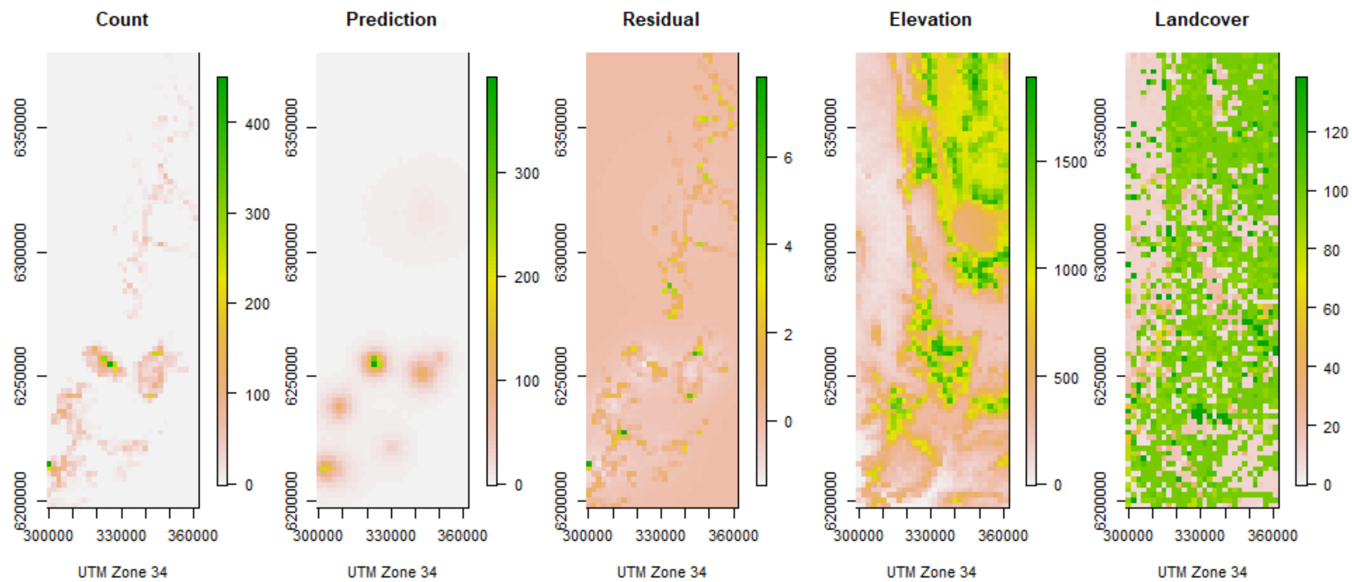

Figure 12

Five pixel plots

These plots demonstrate that most leopards tend to live in areas where the elevation and landcover are in a medium degree. Only one individual prefers to wander in the north area of with high elevation and high landcover, and this individual has a large home range and a small number in pixels based on the predictions of this study. Other individuals have smaller home ranges and high-density space usages in their home ranges. According to Pearson residuals, all individuals in the north and south areas are well predicted by the model; the huge activity range for the individual in the northern area delivers a low density in each pixel. Furthermore, this study finds that there exists a connection between elevation and landcover components and those areas with high elevations tend to have high landcover. In their home ranges, leopards tend to stay in high landcover areas, because such areas could attract more herbivores such as zebra and antelope, thus providing a large amount of food for leopard individuals. In habitat reservation, ecologists should pay more attention to protection of these regions; and the destruction of these areas could cause the biggest effect on leopards.

Discussion

Understanding how animals use their habitats is important for biologists and ecologists. Analysis of landscape components, as explored in this study, could help ecologists to pinpoint animals' habitats and activity areas more efficiently and reduce the habitat destruction, which is nowadays recognized as the main reason for animal extinction (Pimm & Raven, 2000). For instance, 82% of the endangered bird leopard are threatened by habitat loss (Barbault & Sastrapradja, 1995). As one of the most important effort in habitat protection, the habitat identification would let human facilities' building avoid important animal habitats. And understanding the landscape components around animal habitats could provide additional information on the habitats, especially for those animals with large space usage.

Based on the spatial capture-recapture model as well as landscape covariate analysis, this study uses the encounter model to understand each landscape component's effect on animal space usage. The significance of landscape parameters tells whether they influence animal activities in their habitats; the coefficient of parameters shows the components have positive or negative effect on animals' space usage; and the value of the coefficient interprets the degree of effect each component has. Once this study acquires the information about environment elements in the habitats, it could make deeper insights into and protect related areas better and faster.

Leopard analysis

The difference in animal individuals home range could affect the fitness and accuracy of this encounter model. Different leopard individuals have different home ranges, and this causes the elevation and landcover covariates show poor significance in this model. Different home ranges could even affect the relationship between animal count and landscape covariates. This study suggests that the encounter model is fitting with an interaction term of leopard individual category and distance, and this could solve several problems causes by the differences in individual home ranges. When adopting interaction terms, the standard error of all parameters in this model would decrease, compared with the models without interaction terms; and this study builds a model with all significant landscape parameters of elevation and landcover; this means all elevation and landcover elements are associated with leopards' activities. The leopards' space usage is positively associated with elevation as shown in previous models, but when this study uses interaction-term models, the elevation has negative effect on leopard space usage. The landcover coefficient has a positive association with leopard counts in models with or without interaction terms. In the final model, the landcover has larger coefficient than elevation, showing that the landcover component might have stronger effect on leopard than elevation component does. The negative relationship between elevation and space usage indicates that leopards dislike high elevation regions, and the positive relationship between landcover and space usage shows that leopards prefer landcover areas to no-landcover area. Access to prey living in landcover areas and low-elevation areas might be the main reason for such an association, while easy access to rivers or caves could be another reason. In those landcover regions at a relatively low elevation or near a river and lake, there are many herbivores such as zebra and antelope; as a result, these areas attract leopards by their large amounts of food. These areas should be paid more attention to when human facilities are allocated or wild areas are developed. Destruction in these areas could influence leopards seriously, or even decrease leopard individual number.

Extension

We develop the encounter model based on spatial capture-recapture model (Royle 2013b), to analyze landscape components in animal space usage. This model is able to add extension of landscape covariate when estimating individual count in each location based on data. We add interaction term to our model to deal with inequality of individual home range, and change family to quasi-Poisson for huge dispersion parameter. In space usage study, we use our model to estimate individual's count for each pixel, based on the distance of pixel location to individual activity centroid and some additional landscape covariates. To use this model, data should contain distance and landscape information for every pixel. Our model based on assumption that individual has random space usage for each pixel, which means the count in current pixel will not affect the usage in next pixel. This means our model is only sensitive to individual count and density in pixel, but unable to estimate the movement across different pixels. We use this model for independent observations study, and in which doesn't exist individual dependent movement assumption. When the violation of independent assumption occurs, our model might lose huge accuracy and unable to work.

An improvement to this model is adding the movement analysis, especially when individual movement is concerned in the study. Such model could be used for occurrence of those dependent assumptions and could improve estimation of the encounter count by analyzing individual movement based on the spatial capture-recapture model. The encounter model based on ecological distance metrics by Royle (2013 a) is one of the good methods to improve the movement estimation in population study, and his model uses the least-cost path between traps with centroid to replace the Euclidean distance between two points. In this way, the model can work well in space usage study related to animal movement. On the other hand, more landscape covariates added to the model might also be useful to improve the model accuracy.

Conclusion

The model proposed in this study works well in researches containing individual movements. Under the independent activity assumption, this encounter model based on spatial capture-recapture model can be used to analyze the roles of environment elements in animal space usage. The information from the result of this encounter model could help ecologists understand analyze landscape components in animal space usage and solve habitat protection problems, such as those with designating animal sanctuaries. Extension of individual movement with landscape covariates could be added to improve this encounter model when animal space usage is dependent on some specific variables.

References

Primack, R.B. (2006) Essentials of Conservation Biology. 4th Edition. Habitat destruction. Place:177–188. Sinauer Associates, Sunderland, MA.

Hierfl, L.A. (2008). 'Assessing and Prioritizing Ecological Communities for Monitoring in a Regional Habitat Conservation Plan', *Environmental Management*, **42**, 165-179.

Efford, M.G., Dawson, D.K. & Borchers, D.L. (2009) 'Population density estimated from locations of individuals on a passive detector array', *Ecology*, **90**, 2676–2682.

Gardner, B., Royle, J.A., Wegan, M.T., Rainbolt, R.E. & Curtis, P.D. (2010) 'Estimating black bear density using DNA data from hair snares', *The Journal of Wildlife Management*, **74**, 318–325.

Gopalaswamy, A.M., Royle, J.A, Delampady, M, Nichols, J.D, Karanth, K.U. & Macdonald, D.W. (2012) 'Density estimation in tiger populations: combining information for strong inference', *Ecology*, **93**, 1741–1751.

Royle, J.A., Chandler, R.B., Gazenski., K.D. & Graves, T.A. (2013a) 'Spatial capture–recapture models for jointly estimating population density and landscape connectivity', *Ecology,* **94**, 287-294.

Royle, J.A., Chandler, R.B., Gazenski, K.D. & Graves, T.A. (2013b) 'Integrating resource selection information with spatial capture–recapture', *Methods in Ecology and Evolution*, **4**, 520-530.

Royle, J.A. & Nichols, J.D (2003) 'ESTIMATING ABUNDANCE FROM REPEATED PRESENCE–ABSENCE DATA OR POINT COUNTS', *Ecology*, **84**, 777-790.

Efford, M.G. (2004) 'Density estimation in live-trapping studies', Oikos, **106**, 598–610.

Borchers, D.L. & Efford, M.G. (2008) 'Spatially explicit maximum likelihood methods for capture-recapture studies', *Biometrics*, **64**, 377–385.

Royle, J.A. & Young, K.V. (2008) 'A hierarchical model for spatial capture–recapture data', *Ecology*, **89**, 2281–2289.

Krebs, W.N. (1983) 'Ecology of neritic marine diatoms, Arthur Harbor Antarctica', *Micropaleontology*, **29**, 267–297.

Turchin, P. (1998) '*Quantitative analysis of movement. Measuring and modeling population redistribution in plants and animals*', *Sinauer*.

Thompson, W. L., White, G. C. and Gowan, C. (1998) '*Monitoring vertebrate populations*', *Academic Press*.

Dice, L. (1938) 'Some census methods for mammals', *J. Wildl. Manage.* **2**: 119–130.

Anderson, D. R., Burnham, K. P., White, G. C. *et al*. (1983) 'Density estimation of small mammal populations using a trapping web design and distance sampling methods', *Ecology,* **64**, 674–680.

Parmenter, R. R., Yates, T. L., Anderson, D. R. *et al*. (2003) 'Small-mammal density estimation: a field comparison of grid-based vs web-based density estimators' *Ecological Monographs,* **73**, 1–26.

Efford, M.G., Dawson, D.K. & Borchers, D.L. (2009) 'Population density estimated from locations of individuals on a passive detector array', *Ecology*, **90**, 2676–2682.

Borchers, D.L. (2012) 'A non-technical overview of spatially explicit capture–recapture models', *Journal of Ornithology*, **152**, 1–10.

Tischendorf, L. & Fahrig, L. (2000) 'On the usage and measurement of landscape connectivity', *Oikos,* **90**, 7–19.

With, K. & Crist, T. (1995) 'Critical thresholds in species' responses to landscape structure', *Ecology,* **95**, 2446–2459.

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H. & Matthysen, H. (2003) 'The application of 'least-cost' modelling as a functional landscape model', *Landscape and Urban Planning,* **64**, 233–247.

Beier, P., Majka, D. R., & Spencer, W. D. (2008) 'Forks in the road: choices in procedures for designing wildland linkages', *Conservation Biology,* **22**, 836–851.

Zeller, K., McGarigal, K. & Whiteley, A. (2012) 'Estimating landscape resistance to movement: a review', *Landscape Ecology,* 27, 777–797.

Cushman, S. A., McKelvey, K. S., Hayden, J. & Schwartz, M. K. (2006) 'Gene flow in complex landscapes: testing multiple hypotheses with causal modeling', *American Naturalist,* **168**, 486–499.

Schwartz, M. K., Copeland, J. P., Anderson, N. J., Squires, J. R., Inman, R. M., McKelvey, K. S., Pilgrim, K. L., Waits, L. P. & Cushman, S. A. (2009) 'Wolverine gene flow across a narrow climatic niche', *Ecology,* **90**, 3222–3232.

McRae, B. H., & Beier, P. (2007) 'Circuit theory predicts gene flow in plant and animal populations', *Proceedings of the National Academy of Sciences USA,* **104**, 19885.

Pimm, S.L. & Raven, P. (2000) 'Extinction by numbers', *Nature*, 403, 843–845.

Barbault, R. & Sastrapradja, S. (1995) 'Generation, maintenance and loss of biodiversity'. Heywood V.H. Global Biodiversity Assessment. United Nations Environment Program: Cambridge University Press, 193-274.

Chandler, R. B. & Royle, J. A. (2013) ‘SPATIALLY EXPLICIT MODELS FOR INFERENCE ABOUT DENSITY IN UNMARKED OR PARTIALLY MARKED POPULATIONS’. The Annals of Applied Statistics. USGS Patuxent Wildlife Research Center, 936-954.

Appendix

Appendix 1 Additional figures

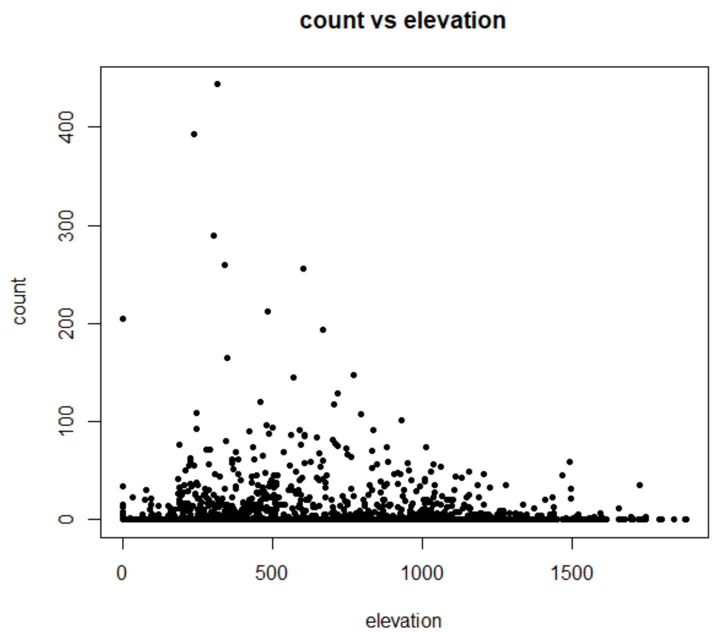

Figure 13 Plot of count and elevation plot based on leopard data

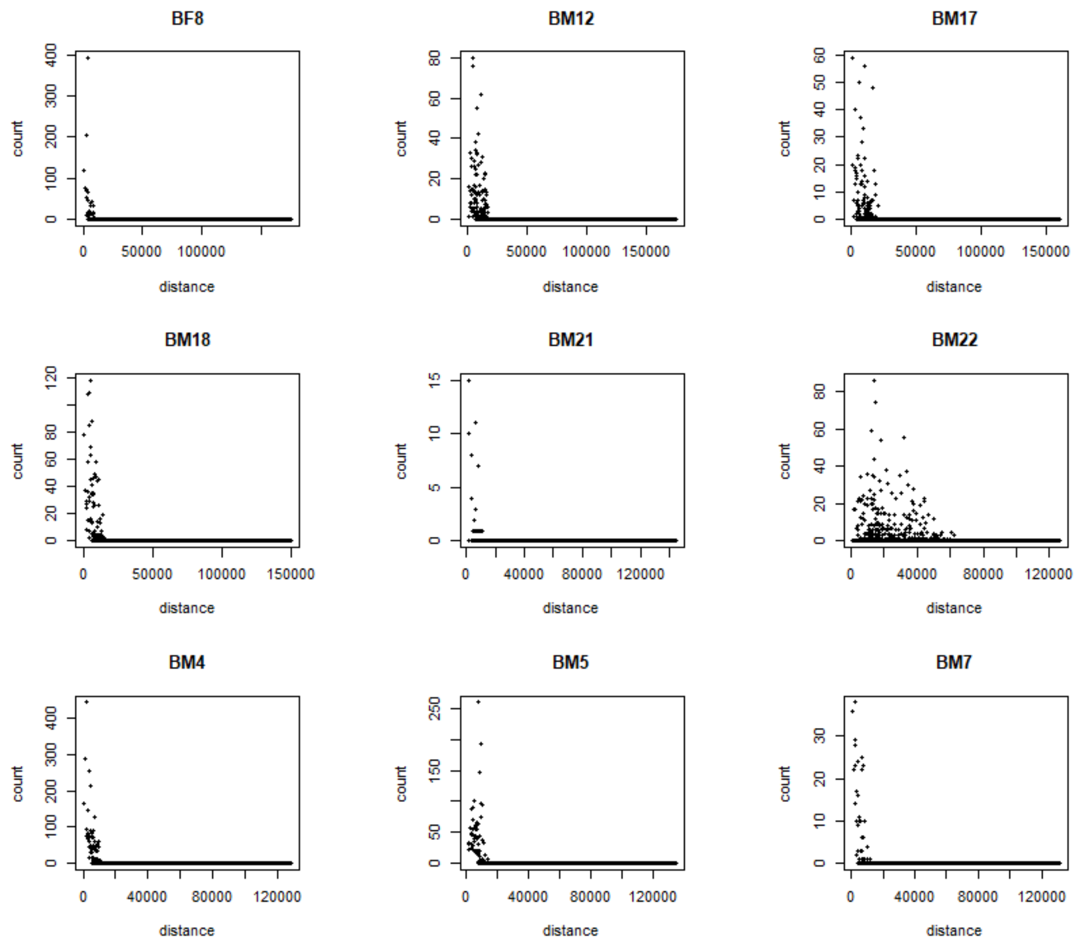

Figure 14 Plot of nine leopard individual distance and count plot

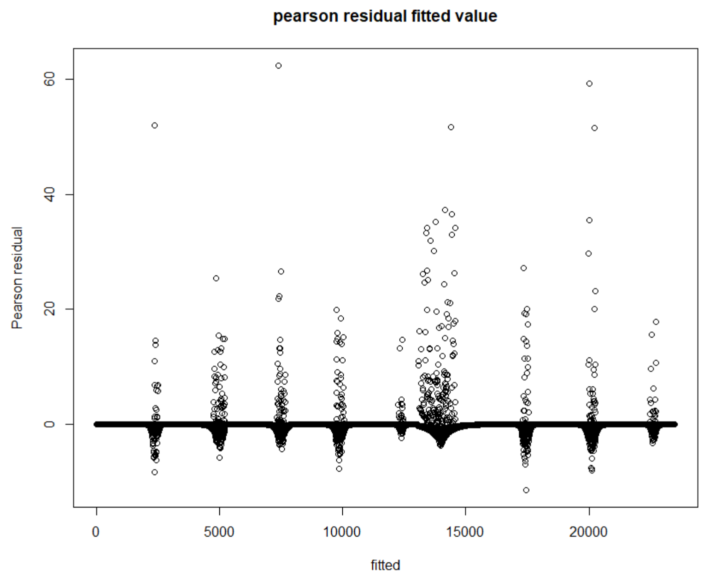

Figure 15 Pearson residuals of fitted value based final model

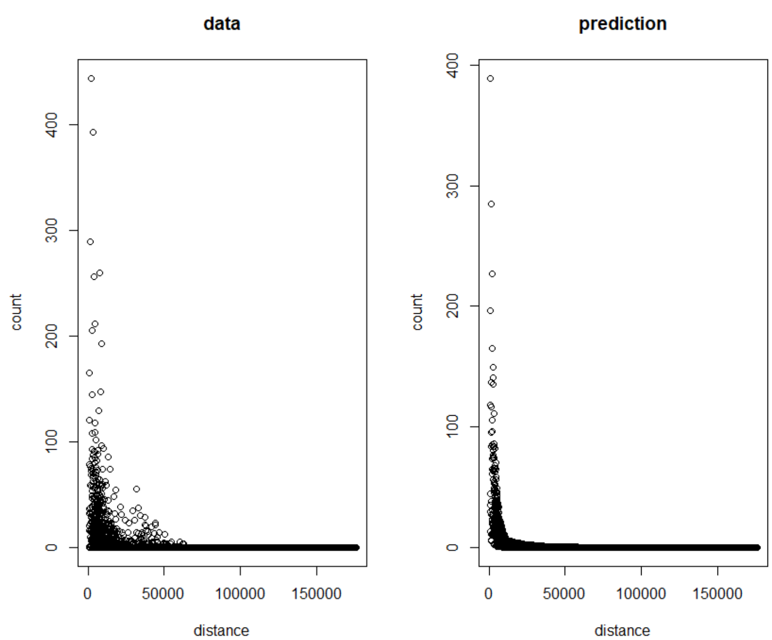

Figure 16 Count vs distance plot of original data and prediction

Appendix 2 Additional tables

Table 3 First 5 rows of data used in final basic encounter model with interaction term. The x and y coordinates are the GPS location of the pixel represented by each row, count is the leopard count, cat the individual name and elevation and landcover are the landscape component covariates.

	x-coordinate	y-coordinate	count	cat	distance	elevation	landcover
1	300173	6378466	0	BF8	165935.2	115	5
2	302273	6378466	0	BF8	165912.6	138	17
3	304373	6378466	0	BF8	165912.6	144	10
4	306473	6378466	0	BF8	165947.1	166	12
5	308573	6378466	0	BF8	166004.3	190	10

Table 4 The QAIC test result for encounter model with interaction term. Basic Model is the encounter model with all terms. DModel is model with only interaction term. EModel is model with interaction term and elevation parameter z_1 (x) and LModel is model with interaction term and landcover parameter z_2 (x).

Model	df	QAIC	loglik	delta	weight
Basic Model	20	11369.23	-15437.80	0	9.999355e-1
DModel	18	11414.61	-15505.09	45.37818	1.400314e-10
EModel	19	11388.53	-15466.82	19.29639	6.453768e-5
LModel	19	11406.59	-15491.44	37.35694	7.727108e-9

Appendix 3 R code

R code used in this paper is uploaded to Github:

https://github.com/xu-zihan/Project_Leopard/blob/master/Data%20Process.R.