Spatial effect on hunter moose (*Alces alces*) observation rates

Supervisor: Lars Rönnegård

Author: Ying Pang, Abdusalaam Abdurasheed

June 10, 2008

Department of Economics and Social Science, Dalarna University

D-level essay in Applied Statistics, 2008

Table of Contents

Abstract

Hunter moose observation rates reported by Swedish hunters are used to monitor moose population density. The accuracy of population density estimates can be improved by finding the suitable model to reflect the relation between hunter moose observation rates and their factor. To improve the accuracy of moose population density, it is necessary to find the important factors affecting the hunter observation rates. Besides hunting team size and area size which is labeled as the essential factors that have most impact on hunter moose observation rates, our main focus is to investigate if spatial effects between hunter moose observations exist or not. Spatial effects include two general aspects which are spatial dependence and spatial heterogeneity. We use the maximum likelihood estimation method and relevant tests based on the standard linear models and spatial models to examine if there are spatial effects. Spatial weight matrix is very important in spatial process models, as the expression of spatial effects. In our case, we believe that the spatial lag dependencies are the main spatial effect in the hunter observation rates, rather than spatial error dependencies. Thus, the spatial autoregressive model achieves the highest validity to fit our data. And we also give suggestions on future research.

Key words: Moose, hunting season, moose hunter observation rate, team size, area size, spatial dependence, spatial weight matrix, spatial models

1. Introduction

Moose is the largest animal of the deer family with long legs and large flat palmate antlers, where they inhabit the forests of northern Asia and Europe from Siberia in the east to Norway in the west, the Baltic region, and northern China. In North America moose are found in wooded areas of Canada and the northern US.

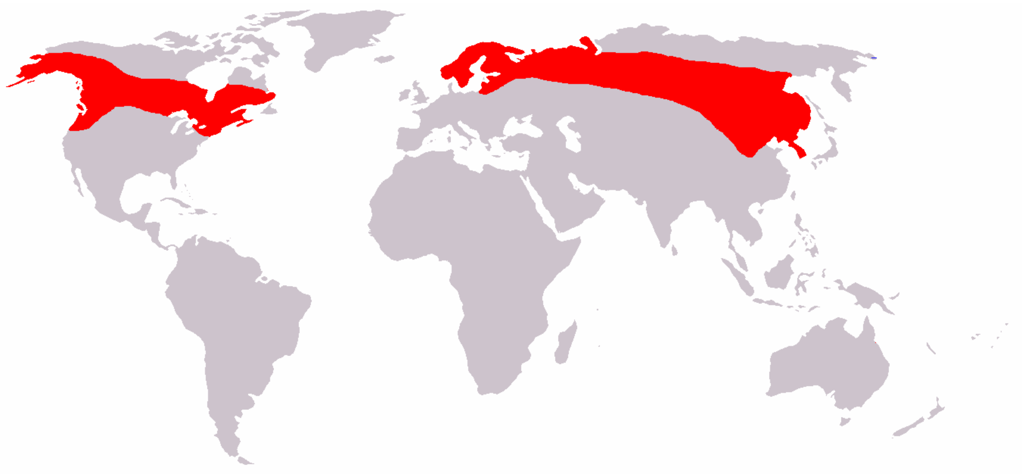

Figure 1.1[1] Moose habitat on the World Map (Red part)

Prominent features of male moose, known as bulls, are the enormous antlers with crown-like shape, which can exceed 1.5 m in width and 22.7 kg in weight and which is why is the reason that moose has been called "the king of the forest" in Sweden. Moose's live weight can range from 350 to 650 kg (**Sand**, 1996). There is also a moderate size dimorphism that males are approximately 15 to 20% larger than females (**Sand**, 1996).

Figure 1.2[2] Young moose in Grönåsens Moose Park, Sweden

[1] Figure 1.1 is taken from WikipediA.
Source: http://en.wikipedia.org/wiki/Image:Moose_distribution.png

The mating system is polygynous with the mating season (rut) taking place from late September to early October. The gestation period lasts for 230-240 days and females normally give birth to one or two calves in late May or early June. Sexual maturity is normally attained at 1.5 or 2.5 years of age, but production of twin calves is normally delayed until 4 years of age. Twinning frequency is highly variable among populations and may range from < 20% to 70%. Triple birth are rare (<0.1%). The female nurses the offspring until late autumn but calves normally remain with their mother until the next parturition (**Sand**, 1996).

With its many lakes and forests, Sweden has one of the highest moose densities in the world (**Ball J.P.** and **Dahlgren J.**, 2002). Therefore, hunting moose is very well integrated in Swedish tradition. Plus, moose meat is nutritious with its low fat content (**Hansson & Malmfors**, 1978; **Hawley, Sylvén & Wilhelmson**, 1983; Crichton, 1998 cited **Sylvén**, 2003). Assuming an average carcass weight of 130 kg (**Hansson & Malmfors**, 1978 cited **Sylvén**, 2003) the moose harvest in 2001 (105,000 moose) contributed with 1,365 million kg of moose meat. This is close up to 10% of the Swedish production of cattle meat in 2001 and more than the Swedish consumption of sheep and horse meat (**Swedish Meat Statistics**, 2001 cited **Sylvén**, 2003). The moose and its harvest are natural and cultural resources which have great economic value (**Hawley, Sylvén & Wilhelmson**, 1983; **Johansson, Kriström & Mattsson**, 1988; **Mattsson**, 1990 cited **Sylvén**, 2003).

In Sweden, hunting season usually lasts from September to December. But in Dalarna County, hunting season starts on the second Monday in October and continues for about 2 months. Moose are hunted from one hour before sunrise until sunset with the allowance of using only rifles. We consider the hunting day to be approximately 8 hours in this article. Hunters in the team share both costs and the meat regardless of which one of them actually shoots the animal.

The hunting rights belong to the landowner who can lease them to another person for shorter or longer periods. Hunting must be conducted in such a manner that the game[3] is not exposed to unnecessary suffering and that people and property are not exposed to danger. About half of the land in Sweden is owned by the state and large companies, particularly in the northern and central regions. On the greater part of this land the hunting rights are leased out to individuals or hunting associations.

The proper management of game species[4] such as moose requires a good monitoring system to determine basic population variables. If there are deficiencies in assessing true population sizes, a seemingly well managed population with sustainable yield and enjoyable recreation components may

---

[2] Figure1.2 is taken from WikipediA.
Source: http://en.wikipedia.org/wiki/Image:Moose-Gustav.jpg

[3] Game is any animal hunted for food or not normally domesticated (from WikipediA).

[4] Hunted animals are referred to as game animals, and are usually large or small mammals, migratory gamebirds, or non-migratory gamebirds (from WikipediA).

in reality be something completely different. Thus, the survey methods used must be well understood, so that the monitoring system can capture the true trends of population change, or even estimate the absolute population size (**Sylvén**, 2003). Several methods including aerial surveys and hunter based surveys have been proposed for detecting trends of population change and absolute sizes (**Seber**, 1982; **Krebs**, 1989; **Caughley & Sinclair**, 1994 cited **Sylvén**, 2003). Although aerial surveys are considered to be the most accurate method for estimating precise moose population sizes and changes, hunter based surveys such as number of moose observed by hunters during hunting season, on the other hand, have been frequently used in Sweden during last decades (**Sylvén**, 2003). Compared to aerial surveys, those based on hunter's observations have lower costs, no dependence on the specific snow conditions required by aerial surveys, less need for adjustment of moose migration effects by counting moose at their hunting ground, and a well anchored hunter involvement in the management of their hunting areas (**Sylvén**, 2003). In Sweden a large number of moose management areas are surveyed annually by the hunting teams. Around 250,000 persons hunt moose each year in Sweden (**Ekman**, 1992 cited **Sylvén**, 2003), and roughly 4.14 million hours were recorded as moose hunting hours by Swedish hunters during the first seven moose-hunting days in 2000 (**K. Kultima**, **Swedish Association for Hunting and Wildlife Management**, **pers.comm.** cited **Sylvén**, 2003).

The hunter moose observation rate, moose seen by a hunter a day, is assumed to reflect the population size (**Ericssson & Wallin,** 1999). Their findings indicate that hunter observation rates of moose reflect moose population size and reproductive rate reasonably well, and can be used to monitor population fluctuations. If calibrated, one may use observation indices for estimates of population size in local moose management as an alternative or supplement to more costly monitoring methods such as aerial surveys. If not calibrated, observational data may be misleading if they are used as a density indicator. **Fryxell et al.** (1988 cited **Ericssson & Wallin,** 1999) also concluded that observed moose per hunter day was a valid index of moose abundance. Similarly, according to **Solberg & Sæther** (1999 cited **Ericssson & Wallin,** 1999) moose observations by hunters accurately reflected fluctuations in moose density and annual reproduction in a moose population in northern Norway. After all, we can use hunter moose observation rates as an indicator of moose population density.

In Scandinavian countries, using the hunter moose observation rate has become essential. There is one article that used hunter moose observation rate to build models with hunting factors including hunting team[5] and team size[6]. **Sylvén** (2003) stated that hunting factors had most impact on the observation rates. Among the observation rates of females, divided in sub-categories according to number of calves in company, only the sub-category females with one calf was significantly affected by

[5] Hunting team refers to all members in an active hunting team collecting moose observation records.
[6] Team size refers to mean number of hunters per examined hunting day in an active hunting team (See section **2**).

the hunting team. Also, team size had significant effects on the sub-category as females with two calves. On the other hand, male observation rates were both affected by hunting team and team size, and female without calves were affected neither of those factors in any of the observation period. For all of these factors, more hunting hours lasts, more hunting team effects. **Sylvén** (2003) also included the area size and evaluated the effect of the size of observation area as in hunting area and moose population density estimates. The result confirms that there is a positive asymptotic relationship between the accuracy of the observation rate and hunting area. As the hunting area increases, so does the accuracy of hunter observation rate. It was concluded that the accuracy of population density estimates could be improved by using hunter observation rates obtained in hunting areas larger than 500 km^2. In our study, we have combined the sub-area all together under their Four-digit area so that we could have better estimates of moose population density (For more details, see section **2**).

In this article, we have included not only the effects of team size and hunting area size towards hunter moose observation rates, but also spatial effects in order to improve the accuracy of the estimates of moose population density. Spatial effects include two general types which are *spatial dependence* and *spatial heterogeneity* (**Anselin**, 1988).

Spatial dependence is generally taken to mean the lack of independence which is often present among observations in cross-sectional data sets[7]. This dependence can be considered to lie at core of the disciplines of regional science and geography, as expressed in **Tobler**'s (1979) *first law of geography*, which states "everything is related to everything else, but near things are more related than distance things." In this sense, spatial dependence is determined by a notion of relative space or relative location, which emphasizes the effect of distance (**Anselin**, 1988).

Spatial heterogeneity is related to the lack of stability over space of behavioral or other relationships under study. More precisely, this implies that functional forms and parameters vary with location and are not homogeneous throughout the data set. For instance, this is likely to occur in econometric models estimated on a cross-sectional data set of dissimilar spatial units, such as rich regions in the north and poor regions in the south. Heterogeneity can be directly related to location in space which is why it is called by the term spatial heterogeneity (**Anselin**, 1988).

Our aim of the essay is to investigate if modeling spatial effects between hunter moose observation rates can improve the accuracy of moose population density. We will examine whether spatial effects exist or not based on a linear regression model. If so, various specifications of spatial models will be performed to modeling spatial effects. And then the spatial model which achieves the highest validity can give the most suitable expression of spatial effects (See section **4**).

---

[7] Cross-sectional data in statistics and econometrics is a type of one-dimensional data set. Cross-sectional data refers to data collected by observing many subjects (such as individuals, firms or countries/regions) at the same point of time, or without regard to differences in time (from WikipediA).

2. Data

In this article, our data is based on two different data sets, which are Hunting Observations provided by **Swedish Association for Hunting and Wildlife Management**[8] and Hunting Areas which is from **Länsstyrelsen Dalarnas Län**[9] (County Administrative Board of Dalarna County).

In Sweden, hunters mostly are members of **Swedish Association for Hunting and Wildlife Management** and **Swedish Hunters Association**[10]. Every year, each hunting event is carried out and recorded by hunting teams with qualified hunters in appointed hunting areas. These two organizations' main objectives are to organize the Swedish hunters, promote good opportunity for hunting, take action on issues of general importance for hunters and wildlife, maintain sustainable numbers of all species through good wildlife management and controlled hunting and work to increase the general understanding of man's responsibility for nature and ecology, while the County Administrative Board decides on different matters concerning responsibilities for issues like informing hunters and landowners, registering hunting grounds, issuing licenses, gathering and compiling shooting statistics at regional level.

2.1 Data set

The data set of Hunting Observations consists of hunting factors of hunting team categories such as Team code, Sub-Area code, Hunting days, ManHours (each Hunting members * Hunting days * 8) and of observation rates of moose population categories such as Adult male, Adult female, Female moose with calves (including Female moose with one calf and Female moose with two calves), Single calf and Unidentified moose and Total (See Table 2.1).

Table 2.1 First few rows from the data set of Hunting Observations

Team code	Sub-Area code	Hunting days	ManHours	Adult male	Adult female	Female with 1 calf	Female with 2 calves	Single calf	Unidentified moose	Total
6201078	1	5	2520	22	25	29	1	1	12	90
6201078	3	5	899	28	36	23	10	1	15	113
6201078	7	5	1529	17	19	32	3	1	16	88
6201078	8	4	435	5	6	6	0	0	1	18

The data set of Hunting Areas consists of Hunting area code (which is the same code as Team code), Type of hunting license, Sub-Area code, Y-coordinate, X-coordinate and Hunting area size (See Table 2.2).

8 Source: http://www.viltdata.se/diana/rptStatistik.aspx?SearchCriteria{20-0-0-0-0-0-0-0-0
9 Source: http://www.lst.se/lst/en/Hunting_and_wildlife_administration.htm
10 Source: http://www.jagareforbundet.se/viltvetande/artpresentation/alg.asp

Table 2.2 First few rows from the data set of Hunting Areas

Hunting area code	Type of hunting license	Sub-Area code	Y-coordinate	X-coordinate	Hunting area size (unit=hectare)
3905035	A	1	6883184	1344894	103.98
3905034	E	3	6889553	1349527	22.87
3905070	B	2	6887336	1318122	7.83
3905009	E	2	6894358	1347354	391.47

2.2 Some pre-treatments of the data

Every data needs some pre-treatments, so does ours. First of all, we have to deal with missing values in the data sets starting by deleting them. Then, we also need to delete manhours with zero value from the data set of Hunting Observations, since we are using observation rate y as a dependence variable in the model.

2.2.1 Combine the data

In Hunting Observation Data, each hunting area code represents a specific hunting area. In each hunting area, there are some sub-areas. For example, in area code 2101002, 2101 represents a bigger area (Four-digit area) and 002 represents its sub-area. The same method goes for the team code, since the Team code of Hunting Observations and the Hunting area code of Hunting Areas are the same as we mentioned above. For the need of building the model from those data sets, we have come to a conclusion that we need to add those sub-areas together under their Four-digit area as well as their hunting observations, since the larger area in which obtaining hunter moose observation rate we are using, the more accurate estimates of moose population density we can get, according to **Sylvén** (2003). Then the data set becomes like this in Table 2.3 as below.

Table 2.3 First few row from the data set of Four-digit areas in Hunting Observations

Team code	Hunting days	ManHours	Adult male	Adult female	Female with 1 calf	Female with 2 calves	Single calf	Unidentified moose	Total
2101	185	17001	110	159	121	38	8	68	504
2103	100	7917	93	90	78	21	7	22	311
2104	64	8155	79	108	91	29	2	30	339
2302	215	45205	684	659	560	118	65	254	2340

As long as we rebuild the data set of Hunting Observations, we should do the same approach to the data set of Hunting Areas. To do so, we need to calculate the weighted average of X-coordinate and Y-coordinate using Area size as the weighting factor and the sum of Area size. The data set is showed below in Table 2.4.

Table 2.4 First few rows from the data set of Four-digit areas in Hunting Areas

Hunting area code	Y-coordinate	X-coordinate	Hunting area size (unit=hectare)
2101	6714135.15	1410847.35	87044.68
2103	6694852.90	1425277.18	50772.24
2104	6699519.34	1400965.57	38636.55
2302	6723003.34	1383761.90	180846.10

Now, we can combine those two data sets of Hunting Observations and Hunting Areas as the model requires. From the data set of Hunting Observations, we only need Hunting days, ManHours and Total. As for Team code, we will prefer using Hunting area code from now on. Here is the result of the combination of two data sets in Table 2.5.

Table 2.5 First few rows from the combination of two data sets

Hunting area code	Hunting days	ManHours	Total	Y-coordinate	X-coordinate	Hunting area size (unit=hectare)
2101	185	17001	504	6714135.15	1410847.35	87044.68
2103	100	7917	311	6694852.90	1425277.18	50772.24
2104	64	8155	339	6699519.34	1400965.57	38636.55
2302	215	45205	2340	6723003.34	1383761.90	180846.10

2.2.2 Variables used in Model

After putting our two data sets together, those variables which are used in the model should be on the table. The dependent variable, observation rate (moose seen per hunter day), which we consider the most in this article, is calculated as

$$y = \frac{\sum N_{Observations}}{\sum N_{ManHours}},$$

where $N_{Observations}$ is number of moose observed from each hunters, $N_{ManHours}$ is the number of manhours for each hunting team, which is equal to $\sum N_{Hunters} \times \sum N_{Days}$, $N_{Hunters}$ is the number of hunters in each team and N_{days} is the length of the observation period (Hunting days) for each hunting team. For the explanatory variables team size and hunting area, we have

$$TeamSize = \frac{\sum N_{ManHours}}{\sum N_{Days} \times 8},$$

and

$$AreaSize = \sum S_{HuntingAreaSize},$$

where $N_{ManHours}$ and N_{Days} are the same that we described above, $S_{HuntingAreaSize}$ is the sum of hunting area size under Four-digit area. Here, we have the final data set combined in Table 2.6 (Also see **Appendix** 1).

Table 2.6 First few rows from the final data set

Area Code	Y-coordinate	X-coordinate	Area size	Days	ManHours	Total	Team Size
2101	6714135.15	1410847.35	87044.68	185	17001	504	11.49
2103	6694852.90	1425277.18	50772.24	100	7917	311	9.90
2104	6699519.34	1400966.57	38636.55	64	8155	339	15.93
2302	6723003.34	1383761.90	180846.10	215	45205	2340	26.28

2.3 Description of Data

Since we don't have a true map of hunting areas, using those X-coordinates, Y-coordinates and Area sizes in our data, we can approximately think of the geographical locations where they collected observations as different circles border with each other or at least neighbor with each other. Figure 2.1 displays the locations. In order to create a first order contiguity matrix[11] of spatial analysis, the circles would provide us an easier and more sensible way.

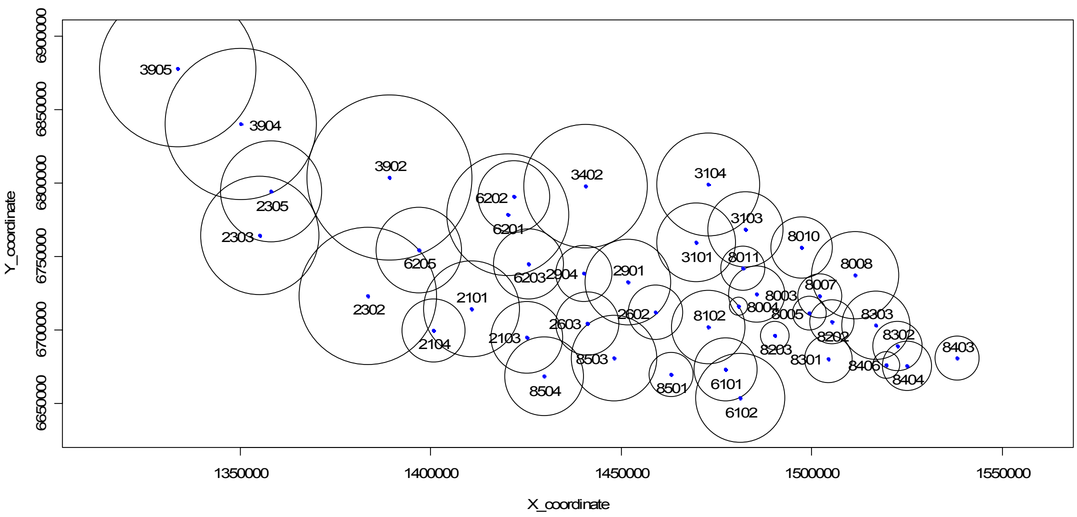

Figure 2.1 Locations of hunter moose observation rates corresponding to their Four-digit area code

In the meanwhile, we can also use hunter moose observation rate in total, y, as in Total in our final data set with their location coordinates such as Y-coordinates and X-coordinates as we used to create location circles to make another type of figure called three-dimensional scatter plot so that we

[11] The more details about first order contiguity matrix will be discussed in **Method** (See section **3**).

could have a first look at observations in Figure 2.2.

From Figure 2.2, we can see that the observation rates are closer to each other at neighboring areas. On the other hand, there are other reasons to believe that there are some spatial correlations exist among observation rates, y. More details will be discussed in **Results** (See section 4).

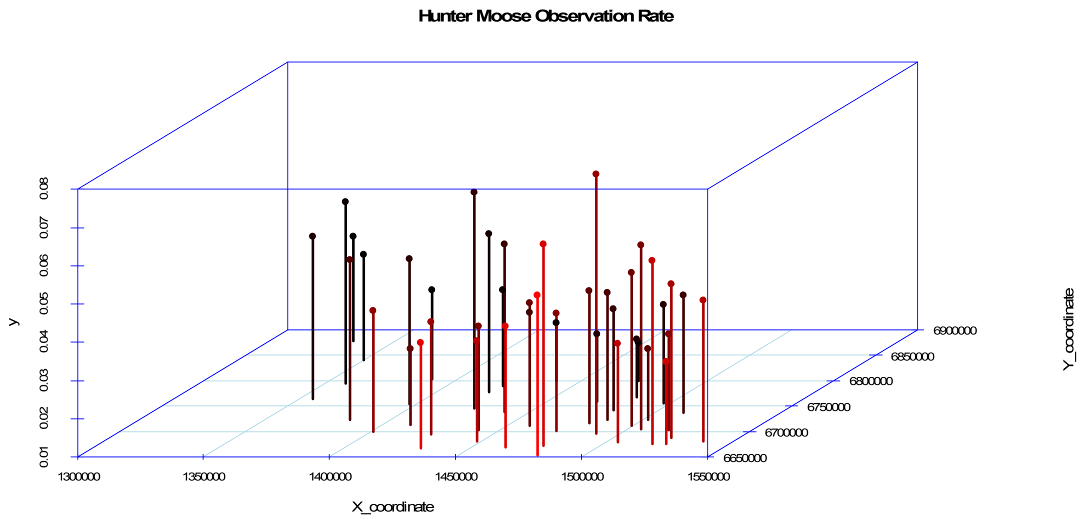

Figure 2.2 Three-dimensional scatter plot for hunter moose observation rate

3. Method

3.1 Linear regression model

Spatial analysis that we use is based on classical linear regression model, so the aim of setting up the linear regression model estimated by ordinary least square (OLS) method is to look for variables that are to be used in spatial models. After choosing variables in section **3**, we have hunter moose observation rate as dependent variable, with considering team size and area size as independent variables.

In linear regression model, we must make sure the dependent variable to meet the statistical assumptions of normality. The Anderson-Darling test is one of the most powerful statistics for assessing whether a sample comes from a normal distribution or not. It may apply to small samples like ours, because the large sample size may reject the null hypothesis of normality with only slight imperfections. With the standard normal cumulative distribution function Φ, the formula for the test statistics is

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left(\ln\Phi(Y_i) + \ln\left(1-\Phi(Y_{n+1-i})\right)\right),$$

the data X_i is standardized as

$$Y_i = \frac{X_i - \overline{X}}{s}.$$

A^{*2}, an approximate adjustment for sample size, is calculated using

$$A^{*2} = A^2\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right).$$

The test statistic can then be compared against the critical values of the theoretical distribution to determine the p-value. The Anderson-Darling test for normality is an empirical distribution function (EDF) test.

Under the hypothesis of normality, data should have kurtosis equal to three and skewness equal to zero. Anscombe-Glynn test of kurtosis and D'Agostino test for skewness have such null hypothesis respectively, so they are quite useful to detect a significant difference of kurtosis and skewness in normally distributed data.

3.2 Spatial model

The main characteristic of spatial modeling is the way in which spatial effects are taken into account. Spatial dependence and spatial heterogeneity, as two aspects to measure the spatial effects, are properly expressed in various spatial modeling situations.

3.2.1 Spatial weight matrix

The original measures for spatial dependence are based on the notion of binary contiguity between spatial units, which are regarded as contiguity or neighbor if they share the same borders (**Moran**, 1948; **Geary**, 1954 cited **Anselin**, 1988). According to this notion, the underlying structure of neighbors is expressed by 0-1 values. If two spatial units have a common border of non-zero length they are considered to be contiguous, and a value of 1 is assigned.

The simple concept of binary contiguity was extended by **Cliff** and **Ord** (1973, 1981 cited **Anselin**, 1988) to include a general measure of the potential interaction between two spatial units. This is expressed in a spatial weight matrix $\mathbf{W}$, also referred to as a Cliff-Ord weight matrix. Spatial weight matrix, as a typical way which makes space has been formalized, is used to achieve that spatial models can be applied to many empirical contexts.

However, the spatial linkages cannot always be properly defined by using the geographical information in application. In our case, we couldn't get a true map of hunting areas which shows us common borders with positive length. Thus, we have come to a solution that we should use Figure 2.1 to define which areas are neighbors, and then construct the spatial weight matrix.

Since observation rates are available for a cross-section of spatial units, at one point in time, we focus on the first order contiguity. But the weight matrix used in spatial modeling is usually row-standardized, scaled such that the sum of the row elements is equal to one, to make sure its asymmetry.

3.2.2 Spatial models

The taxonomy of spatial models is obtained by assuming homoscedasticity, because we prefer to focus primarily on the specification of the spatial dependence.

A general version of the spatial model, called SAC model, includes both the spatially lag term and a spatially correlated error structure. It is written as follows

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\rho}\mathbf{W}_1\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \boldsymbol{\lambda}\mathbf{W}_2\boldsymbol{\varepsilon} + \boldsymbol{\mu} \\ \boldsymbol{\mu} &\in N(0, \sigma_\mu^2 \mathbf{I}_\mathbf{n}) \end{aligned} \tag{3.1}$$

Where,

$\mathbf{y}$ represents a N by 1 vector of observations of dependent variable including n locations,

$\mathbf{X}$ is a N by K matrix of exogenous variables,

$\boldsymbol{\beta}$ is a K by 1 vector of parameters associated with exogenous variables,

ρ is the coefficient on the spatially lagged dependent variable $\mathbf{y}$,

λ is the coefficient in the spatial autoregressive structure for the disturbance ε,

$\mathbf{W}_1$ and $\mathbf{W}_2$ are two N by N spatial weight matrices, respectively associated with a spatial autoregressive process in the dependent variable and in the disturbance term. In order to be identified, it is necessary that the weight matrices for spatial autoregressive terms in the dependent variable and in the errors are different, namely $\mathbf{W}_1 \neq \mathbf{W}_2$.

If a model is referred to as the SAC model, it must be satisfied with both $\rho \neq 0$ and $\lambda \neq 0$. Moreover, there are two familiar special cased for the SAC model. One is to set $\lambda = 0$, the SAC model is then expressed as follows

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\rho}\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &\in N(0, \sigma_\mu^2 \mathbf{I}_\mathbf{n}) \end{aligned} \tag{3.2}$$

where $\mathbf{W}$ is spatial weight matrix, and all others are as same as those defined in the SAC model.

Model (3.2) is referred to as the spatial autoregressive model (SAR model), which is also labeled as the standard simple regression model with a spatially lagged dependent variable. In other words, it is a special version of the spatial model which only includes the spatial dependence of spatial lagged term. **Anselin** (1988) provides a maximum likelihood (ML) method to estimate the parameters for this model. The steps of ML estimation for SAR model are listed in **Anselin** (1988) as follows:

i. Carry out ordinary least squares (OLS) for the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{\mathbf{Y}} + \boldsymbol{\varepsilon}_{\mathbf{Y}}$.

ii. Carry out OLS for the model $\mathbf{Wy} = \mathbf{Xb}_{\mathbf{W}} + \boldsymbol{\varepsilon}_{\mathbf{W}}$.

iii. Calculate residuals $\mathbf{e}_{\mathbf{Y}}\left(= \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{Y}}\right)$ and $\mathbf{e}_{\mathbf{W}}\left(= \mathbf{Wy} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{W}}\right)$.

iv. Based on $\mathbf{e}_{\mathbf{Y}}$ and $\mathbf{e}_{\mathbf{W}}$, estimate ρ that maximizes the concentrated likelihood function $L_C = C - \frac{n}{2}\ln\left[\frac{1}{n}\left(e_Y - \rho e_W\right)'\left(e_Y - \rho e_W\right)\right] + \ln\left|I - \rho W\right|$, where C is a constant.

v. Based on the $\hat{\rho}$ that maximized L_C, calculate $\hat{\boldsymbol{\beta}}\left(= \hat{\boldsymbol{\beta}}_{\mathbf{Y}} - \boldsymbol{\rho}\hat{\boldsymbol{\beta}}_{\mathbf{W}}\right)$ and $\hat{\sigma}_{\varepsilon}^2\left(= \frac{1}{n}\left(e_Y - \rho e_W\right)'\left(e_Y - \rho e_W\right)\right)$.

The other version of SAC model sets $\rho=0$ and it is written as follows

$$\begin{aligned} &\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ &\boldsymbol{\varepsilon} = \boldsymbol{\lambda}\mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\mu} \\ &\boldsymbol{\mu} \in N(0, \sigma_{\mu}^2\mathbf{I}_{\mathbf{n}}) \end{aligned} \tag{3.3}$$

where $\mathbf{W}$ is spatial weight matrix, and all others are as same as those defined in the SAC model.

Model (3.3) is the spatial error model (SEM model), which the disturbances display spatial dependence. It is also regarded as a linear regression model with a spatial autoregressive disturbance. The ML estimation for parameters in the SEM model is relatively more complex and is performed by the following steps according to **Anselin** (1988):

i. Perform OLS for the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{\mathbf{OLS}} + \boldsymbol{\varepsilon}_{\mathbf{OLS}}$.

ii. Calculate residuals $\mathbf{e}_{\mathbf{OLS}}\left(= \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right)$.

iii. Based on $\mathbf{e}_{\mathbf{OLS}}$, estimate λ that maximizes the concentrated likelihood function $L_C = C - \frac{n}{2}\ln\left(\frac{1}{n}e_{OLS}'B'e_{OLS}\right) + \ln\left|I - \lambda W\right|$, where C is a constant, $\mathbf{B} = \mathbf{I} - \boldsymbol{\lambda}\mathbf{W}$.

iv. Based on the $\hat{\lambda}$ that maximized L_C, perform Estimated Generalized Least Squares for the model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{\mathbf{EGLS}} + \boldsymbol{\varepsilon}_{\mathbf{EGLS}}$.

v. Calculate residuals $\mathbf{e}\left(=\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{EGLS}}\right)$

vi. If the convergence criterion is met, proceed to step vii, otherwise return to step iii.

vii. Based on $\mathbf{e}$ and λ, calculate $\hat{\sigma}^2\left(=\frac{1}{n}e'B'e\right)$, where $\mathbf{B}=\mathbf{I}-\boldsymbol{\lambda}\mathbf{W}$.

In analogy to the well known Cochrane-Orcutt (1949) and Durbin (1960) procedures developed for the case with serial error correlation in the time series, similar approaches have been suggested for the spatial model. The spatial Durbin approach is based on the formal equivalence of the model to a spatial autoregressive specification. A spatial error model can also be specified in spatial lag form, with the spatially lagged explanatory variables included. The spatial lag form of the error model is referred to as the spatial Durbin model. It is shown as follows:

$$\begin{aligned}&\mathbf{y}=\boldsymbol{\lambda}\mathbf{W}\mathbf{y}+\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\lambda}\mathbf{W}\boldsymbol{\beta}+\boldsymbol{\mu}\\&\boldsymbol{\mu}\in N(0,\sigma_{\mu}^{2}\mathbf{I}_{\mathbf{n}})\end{aligned} \tag{3.4}$$

The starting point is a spatial error model

$$\begin{aligned}&\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}\\&\boldsymbol{\varepsilon}=\boldsymbol{\lambda}\mathbf{W}\boldsymbol{\varepsilon}+\boldsymbol{\mu}\end{aligned}$$

or

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\left(\mathbf{I}-\boldsymbol{\lambda}\mathbf{W}\right)^{-1}\boldsymbol{\mu}.$$

After pre-multiplying both sides of this equation by $\left(\mathbf{I}-\boldsymbol{\lambda}\mathbf{W}\right)$, a spatial autoregressive model results, as in the spatial Durbin approach:

$$\mathbf{y}=\boldsymbol{\lambda}\mathbf{W}\mathbf{y}+\mathbf{X}\boldsymbol{\beta}-\mathbf{X}\boldsymbol{\lambda}\mathbf{W}\boldsymbol{\beta}+\boldsymbol{\mu}$$

or

$$\mathbf{y}=\boldsymbol{\lambda}\mathbf{W}\mathbf{y}+\mathbf{X}\boldsymbol{\beta}+\mathbf{W}\mathbf{X}_{\boldsymbol{\gamma}}+\boldsymbol{\mu}$$

Reversing the flow of reasoning, the more complex autoregressive model can be reduced to a simple of $\mathbf{y}$ on $\mathbf{X}$ with a spatially autoregressive disturbance. More precisely, in order for the two specifications, (3.3) and (3.4), to be equivalent, the product of the coefficients of $\mathbf{Wy}$ and $\mathbf{X}$ (λ times β) should be equal the negative of the coefficient of $\mathbf{WX}$ ($\lambda\cdot\beta$ or $-\gamma$).

A maximum likelihood estimation model (3.4) will yield consistent estimates for λ and β. Because of including the spatial lag term and spatial error term, spatial Durbin model is very useful when the SAC model can not identified.

3.2.3 Hypothesis testing

Given the widespread use of the maximum likelihood method in the estimation of spatial process models, most hypotheses for the parameters of these models are based on asymptotic considerations as well. Most of the inference in spatial model is based on the Wald test, or, equivalently, the asymptotic t-test and Likelihood Ratio (LR) test.

In contrast to the Wald and LR approaches, tests based on the Lagrange Multiplier (LM) principle do not necessitate the estimation of the more complex model, just linear model with OLS estimation in our case. As shown in **Anselin** (1988), a formal derivation of the LM test for the null hypothesis of presence of spatial dependence yields a complex statistic.

The test that pertains to spatial dependence does not have a direct counterpart. However, the resulting expression is still given as follows:

$$D^{-1} \cdot \left(R_y - R_e\right)^2 + \left(1/T\right) R_e^2,$$

where

$$\mathbf{R_y} = \mathbf{e}'\mathbf{W} \cdot \mathbf{y} / \sigma^2,$$

$$\mathbf{R_e} = \mathbf{e}'\mathbf{W} \cdot \mathbf{e} / \sigma^2,$$

$$\mathbf{M} = \mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}',$$

$$\mathrm{T} = \mathrm{tr}\left\{\left(\mathbf{W}' + \mathbf{W}\right)\mathbf{W}\right\},$$

and

$$D = \sigma^{-2}\left(\mathbf{WX\beta}\right)' \mathbf{M}\left(\mathbf{WX\beta}\right),$$

$\mathbf{e}$ is a vector of OLS residuals, and others are as same as those defined above.

It consists of two parts, one pertaining to the spatial lag autocorrelation component, the other to the spatial error autocorrelation component. Thus, LM test can form the basis for a wide range of tests, one-directional as well as multidirectional.

$\mathbf{R_e}$ is relative to the Moran test, which is a traditional test especially for the presence of spatial autocorrelation in the error term. A Moran I statistic can be applied to regression residuals in a straightforward way. The expression is

$$I = \mathbf{e}'\mathbf{W}\mathbf{e}/\mathbf{e}'\mathbf{e} = \mathbf{e}'\mathbf{W}\mathbf{e}/\sigma^2 .$$

Besides, the Breusch-Pagan test is usually to test for heteroskedasticity in a linear regression model, and the spatially adjusted Breusch-Pagan test is widely used to test for the presence of spatial heteroskedasticity in spatial models. The test statistic is given:

$$\left(\frac{1}{2}\right)\mathbf{f'X}\left(\mathbf{X'X}\right)^{-1}\mathbf{X'f} \sim \chi^2\left(K-1\right),$$

with $f_i = \sigma^{-2} e_i - 1$ in which e_i is the OLS residual for observation i.

Anselin (1988) provides a maximum likelihood (ML) method to estimate the parameters for the SAR model and SEM model. The 'spdep' package in R software can create the neighbor list and weight matrix, perform ML estimations for spatial models, including spatial autoregressive model, spatial error model and spatial Durbin model, and show the relative test results.

3.3 Model selection

The use of information theory as a tool in model building and model validation in regional science and geography has been advocated from a number of different viewpoints.

In parametric methods, there might be various candidate models with different number of parameters. The likelihood function is increased when the number of parameters in the model is increased but it might result in overtraining problem if the number of parameters is too large. In order to overcome this problem one can use Schwarz Criterion (SC) which is a statistical criterion for model selection. The formula for the SC is

$$SC = -2\ln(L) + k\ln(n),$$

where k is the number of parameters in the model, n is the number of observations, and L is the maximized value of the likelihood function for the estimated model. The model with the lower value of SC is the one to be preferred, because lower SC implies either fewer explanatory variables or better fit, or both.

Akaike's information criterion (AIC) is a measure of the goodness of fit of an estimated model. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best. In general terms, this measure can be formally expressed as:

$$AIC = 2k - \ln(L).$$

The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters. It penalized free parameters less strongly than SC.

4. Results

4.1 Linear model

We need to test for normality of response variable at first. The relative test for normality of the observation rate, y, is given in the Table 4.1. The p-values are high, which means that we cannot reject the null hypothesis that the data comes from a normal distribution. A significance level of 10% is used.

Table 4.1 Results of Test for Normality

	Statistic	P-value
Anderson-Darling normality test	0.4909	0.2084
Anscombe-Glynn kurtosis test	3.4050	0.3214
D'Agostino skewness test	0.6431	0.2306

The use of ordinary least squares to the linear regression model of the observation rate on team size, area size and their interaction values yields the following result, with the estimated standard error and corresponding t-value listed in parentheses:

$$y = 2.201\times10^{-3}\left(TeamSize\right)+3.173\times10^{-7}\left(AreaSize\right)\text{-}1.787\times10^{-8}\left(TeamSize*AreaSize\right)$$

$$(3.098\times10^{-4}) \qquad (9.129\times10^{-8}) \qquad (6.835\times10^{-9})$$

$$(7.125) \qquad (3.475) \qquad (-2.614)$$

$$R^2 = 0.7647 \qquad R^2_{adj.} = 0.7647$$

All estimated coefficients are strongly significant. Moreover, the regression achieves a reasonable fit in terms of R-square and adjusted R-square value of more than 0.7.

4.2 Spatial model

4.2.1 Spatial weight matrix

Spatial dependence is expressed by a spatial weight matrix and taken as first order contiguity between neighborhoods. According to Figure 2.1, the first order contiguity is illustrated in Table 4.2 (See **Appendix** 2)

Table 4.2 First few rows of the first order contiguity, hunting areas, Dalarna Neighborhoods

Neighborhood	Contiguous to:					
2101:	2103	2104	2302	6201	6203	6205
2103:	2101	2603	8504			
2104:	2101	2302				
2302:	2101	2104	2303	3902	6205	
2303:	2302	2305	3904			

According to the first order contiguity, spatial weight matrix, $\mathbf{W}$, is created. The elements of the matrix, w_{ij}, are assigned the value of 1 if hunting area i and j are contiguous. Otherwise, the value of w_{ij} is 0. Then, the row-standardized weight matrix is used in the tests and models.

4.2.2 Tests and estimations of spatial models

Based on the linear regression model above, an LM test is performed to test for the presence of spatial dependence. The results for the various one-directional and multidirectional tests are listed in Table 4.3.

Table 4.3 Lagrange Multiplier Diagnostics for Spatial Dependence in OLS Linear Regression

	H_0	H_1	χ^2	df	p-value
One Direction Tests					
Spatial Error Autocorrelation	$\lambda = 0$ and $\rho = 0$	$\lambda \neq 0$ and $\rho = 0$	1.678	1	0.1953
Spatial Lag Dependence	$\rho = 0$ and $\lambda = 0$	$\rho \neq 0$ and $\lambda = 0$	9.137	1	0.0025
Multidirectional Tests					
Spatial Error Autocorrelation	$\lambda = 0$	$\lambda \neq 0$	3.108	1	0.0779
Spatial Lag Dependence	$\rho = 0$	$\rho \neq 0$	10.567	1	0.0012
All Spatial Effects	$\lambda = 0$ and $\rho = 0$	$\lambda \neq 0$ and $\rho \neq 0$	12.245	2	0.0022

Clearly, the highest significance is achieved for a test against the absence of spatial lag dependence. Thus, we will estimate SAR model, adding the spatial lag dependence on the linear regression. The ML estimate of spatial autoregressive model is given in Table 4.4, with corresponding t-values listed in parentheses.

The traditional test for the presence of the spatial autocorrelation in the error term is based on the use of the Moran I statistic computed for the residuals, as mentioned in method section. For the linear regression above, this statistic is 0.1571, highly significant with a probability of 0.0454. We should reject the hypothesis of the absence of spatial correlation in error terms. Furthermore, the LM test for spatial error dependence in multi-direction has a slightly lower significance than that indicate by a traditional Moran test, although they are shown the same result of the presence of spatial error autocorrelation. Therefore, it is necessary to estimate SEM model. The result of spatial error model based on ML approach and corresponding t-value is listed in Table 4.4.

And the estimation of the linear regression model is also shown in Table 4.4 to compare the models with spatial dependence or without it.

Table 4.4 ML Estimations in the Spatial Autoregressive Model and Spatial Error Model

Variables	Dependent variable: Y		
	SAR	SEM	Linear
ρ	0.53843 (4.7275)	-	-
λ	-	0.41054 (2.638)	-
TeamSize	1.152×10^{-3} (3.608)	1.822×10^{-3} (5.366)	2.201×10^{-3} (7.125)
AreaSize	1.295×10^{-7} (1.612)	2.866×10^{-7} (2.832)	3.173×10^{-7} (3.475)
TeamSize*AreaSize	-8.709×10^{-9} (-1.536)	-1.575×10^{-8} (-2.189)	-1.787×10^{-8} (-2.614)

All the estimated coefficients of explanatory variables are significant according to the asymptotic t-test, both in the SAR model and SEM model. However, the interpretations for the coefficients of the explanatory variables are not our main focus.

Under the ML approach, the high significance is achieved for testing the coefficient of spatial lag variable, ρ, both in LR test and Wald test.

An asymptotic t-test or Wald test clearly indicates the significance of spatial parameter in spatial error model, λ, and the corresponding values are 2.6378 and 6.958. However, the resulting statistic of LR test is 2.5267, with a probability level of 0.11193 which is not significant enough. Different tests give opposite significance of spatial parameter, which is as same as the LM test of spatial error autocorrelation from one-direction and multi-direction. Thus, we have cause to suspect the presence of spatial error dependence, and it is necessary to estimate the model including both spatial autoregressive dependence and spatial error dependence.

As pointed out in the method section, an alternative approach to estimation in a model with both spatial lag and error autocorrelation consists of the spatial Durbin method. The resulting estimates based on ML method are presented in Table 4.5, with corresponding t-value.

Table 4.5 Estimation in the Spatial Durbin Model

Variables	estimate	t-value
λ	0.38121	2.4454
TeamSize	6.5317×10^{-4}	1.6659
AreaSize	-1.6005×10^{-8}	-0.1235
TeamSize*AreaSize	-3.2597×10^{-9}	-0.4262
Lag.(TeamSize)	9.6614×10^{-4}	1.5637
Lag.(AreaSize)	2.1038×10^{-7}	1.2572
Lag.(TeamSize*AreaSize)	-9.3956×10^{-9}	-0.8807

Whereas the estimated λ parameter is clearly significant, giving the p-value of 0.0331, however the coefficients of the lagged explanatory variables are not. Moreover, for Lag.(TeamSize) and Lag.(TeamSize*AreaSize), the wrong sign is obtained, since the product of λ and coefficients of explanatory variables should be equal the negative of the coefficient of lagged explanatory variables as proved in method section.

4.3 Model comparison

The log-likelihood and corresponding AIC and SC are listed in Table 4.6, where for four of the estimated models above.

Table4.6 Information Based Measures of Fit for Spatial Models (Ranking in Parentheses)

Model	L	Rank	AIC	Rank	SC	Rank
Linear Regression	101.73	4	-195.45	4	-188.50	4
SAR	108.00	2	-205.99	1	-207.88	1
SEM	103.00	3	-195.98	3	-197.86	3
Spatial Durbin	110.19	1	-204.37	2	-207.38	2

The ranking of models implied by both information measured of fit is the same, although the relative magnitudes for the AIC and SC are clearly different. Also, the trade-off between fit and parsimony, which is encompassed in the AIC and SC, yields a different indication of model validity than the simple log-likelihood. According to the latter, as in Table 4.6, the spatial Durbin model is best. However, since the AIC and SC take into account the relative lack of additional fit provided by the three extra explanatory variables in this model, its corresponding ranking is affected, and it is shown to be worse than the spatial autoregressive model. In addition, we perform the Moran test and adjusted Breusch-Pagan test for the spatial autoregressive model. Given the Moran I statistic is -0.0594 with corresponding p-value of 0.6092 and B-P test statistic is 1.8653 with a probability of 0.3935, there are sufficient evidence to accept the hypotheses of absence of spatial error autocorrelation and spatial heteroskedasticity. Overall, the spatial autoregressive model achieves the highest validity.

5. Conclusion

According to the analysis and results above, the spatial autoregressive model is chosen to express spatial effects in our data, which provides sufficient evidence that the spatial lag dependence is the main spatial effect, rather than spatial error autocorrelation.

$$y = 0.53843(\mathbf{W}Y) + 1.152\times10^{-3}(TeamSize) + 1.295\times10^{-7}(AreaSize) - 8.709\times10^{-9}(TeamSize * AreaSize)$$

Team size and area size have the positive influence on the observation rates, and meanwhile the sign of the interaction terms is negative. The product term may further complicate interpretation. If we focus on the influence of team size, fixed area size, it is easy to include that the hunter moose observation rate always increase along with the increase of team size. However, when fixed team size, we can not believe that more observations will occurs in the lager area. It is also explained from the actual aspect. Assuming that two hunting teams have the same team size, the hunting team with the larger hunting area may have less observation rates whatever the true moose population density is, since they don't have enough team members to cover the whole area. On the contrary, the one with smaller hunting area can be able to cover the whole area so that their observations would be closer to the actual moose population.

The lagged dependent variable, as a new explanatory variable, shows spatial autoregressive dependence on hunter moose observation rates. The significance of its coefficient estimate indicates the necessity of taking spatial effects into account, which means ignoring spatial contiguity information of the data will lead to an inappropriate model specification. In other words, modeling the spatial effects between hunter moose observation rates can better fit our data, which also can improve the accuracy of the moose population density.

6. Discussion

6.1 About the sample size

Since the properties of estimators and tests for spatial models are based on asymptotic considerations, their application in small samples may be problematic. The asymptotic equivalence of the Wald, Likelihood Ratio and Lagrange Multiplier approaches fails to be reflected in finite samples and may lead to conflicting indications.

Much remains to be done in order to obtain better insight into the importance of sample size. One possible avenue for research would consist of the further development of finite sample approximations. Alternatively, inference could be based on Bayesian strategies or on resampling procedures such as the bootstrap. In addition, a much large body of simulation results is needed to provide better guidance in the choice of estimators and tests for spatial models.

6.2 About choice of weight matrix

An important methodological problem in the estimation of spatial process models is associated with the proper choice of the weights. A misspecified weight matrix may result in inconsistent estimates and misleading inference. Moreover, the uncertainty associated with the value of the weights, which is typically ignored, may affect the interpretation of the estimation results.

The matter follows from the lack of uniform criteria for the choice of a spatial weight matrix. Competing weight matrices can often be considered for the same empirical context, with no clear a priori criteria to guide the choice. Since the properties of estimators and tests will be determined in part by the form of spatial dependence incorporated in these weights, it is important that the matrices can be characterized in formal terms. Consequently, it is crucial to be able to assess the validity of the assumptions embedded in the weight matrix.

6.3 About spatial correlation structures

In our case, we focus on spatial models, which spatial dependence is expressed by spatial weight matrix. Actually, there are other methods that can realize modeling spatial effects, and one of them is spatial correlation structures. They were originally proposed to model dependence in the data indexed by continuous two-dimensional position vectors. Because the spatial correlation structures are continuous functions of some distance between position dimensions, they are easily generalized to any finite number of position dimensions. The basic reference for spatial correlation structures used with linear models with no random effects is **Cressie** (1993 cited **José C. Pinheiro** and **Douglas M. Bates**, 2000). Spatial correlation structures in the context of mixed-effects models are described at length in **Diggle et al.** (1994 cited **José C. Pinheiro** and **Douglas M. Bates**, 2000). If more information about original hunting team and their hunting area can be collected, then it is realized to model dependence in grouped-data with spatial correlation structures by mixed-effects models.

Spatial correlation structures are generally represented by their semivariogram, instead of their correlation function (**Cressie**, 1993, §2.3.1). The semivariogram of a spatial correlation structure is a function with distance between locations, which any distance metric may be used. If there is no spatial correlation between locations, then the semivariogram will be constant. Semivariogram can be expressed by observations $z_i, i = 1,...,k$ at locations $x_1,...,x_k$ as

$$\hat{\gamma}(h) \approx \frac{1}{2\,|\,N(h)\,|} \sum_{(i,j)\in N(h)} \left| z_i - z_j \right|^2,$$

where $N(h)$ denotes the set of pairs of observation i, j placed at an approximate distance of h and $|N(h)|$ is the number of pairs in the set (**Cressie**, 1993).

Given the X and Y coordinate of hunting areas, we try to show the plot of spatial semivariogram (using residuals $\hat{\varepsilon}_i$ as observations z_i) versus distance between areas based on the linear model and the spatial lag model, respectively in Figure 6.1(a) and Figure 6.1(b).

Figure 6.1(a)

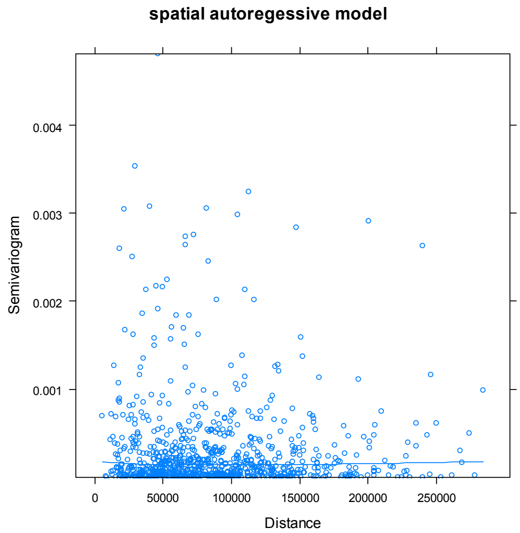

Figure 6.1(b)

Figure 6.1 cloud plots of semivariogram versus distance calculated by residuals of linear regression model and spatial autoregressive model

Since there is no concern of spatial effects in linear regression model, the semivariogram increases monotonically with distance depending on spatial correlation among residuals. On the contrary, considering spatial effects in spatial autoregressive model shows smooth flat line of semivariogram with no sign of spatial correlation among the residuals. And if more information of hunting areas can be collected, the semivariogram models for spatial correlation structure may realize based on the data grouped by locations. We believe this will be a nice direction for future research on this topic.

Reference

1. **Sand, Håkan**, 1996. Life history strategies in moose (*Alces alces*): geographical and temporal variation in body growth and reproduction. Doctoral dissertation, Uppsala. ISBN 91-576-5116-7.
2. **Sylvén, Susanne**, 2003. Management and Regulated Harvest of Moose (*Alces alces*) in Sweden. Doctoral dissertation, Uppsala. ISBN 91-576-6402-1
3. **Ericsson, G. & Wallin, K.**, 1999. Hunter observations as index of moose *Alces alces* population parameters. – Wildl. Biol. 5: 177-185
4. **Ball J.P., Dahlgren J.**, 2002. Browsing Damage on Pine (Pinus sylvestris and P. contorta) by a migrating moose (*Alces alces*) Population in Winter: Relation to Habitat Composition and Road Barriers. Scandinavian Journal of Forest Research, 17 (5), 427-435(9)
5. **Stephens, M.A.**, (1986). Tests based on EDF statistics. In: **D'Agostino, R.B.** and **Stephens, M.A.**, eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.
6. **Thode Jr., H.C.** (2002): Testing for Normality. Marcel Dekker, New York
7. **Anscombe, F.J., Glynn, W.J.** (1983) Distribution of kurtosis statistic for normal statistics. Biometrika, 70, 1, 227-234
8. **D'Agostino, R.B.** (1970). Transformation to Normality of the Null Distribution of G1. Biometrika, 57, 3, 679-681
9. **Anselin, Luc.** 1988 Spatial econometrics: methods and models. (Dordrecht: Kluwer)
10. **Anselin, Luc., Bera, A. K., Florax, R.** and **Yoon, M. J.** 1996 Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, 26, 77–104
11. **Sakamoto, Y., Ishiguro, M., and Kitagawa G.** (1986). Akaike Information Criterion Statistics. D. Reidel Publishing Company
12. **José C. Pinheiro** and **Douglas M. Bates** (2000). Mixed-Effects Models in S and S-Plus. Springer-Verlag, New York
13. **Schwarz, G.** (1978) "Estimating the Dimension of a Model", Annals of Statistics, 6, 461-464
14. **Cressie,N.**, 1993, Statistics for spatial data, Wiley Interscience

Appendix 1

Area Code	Y_coordinate	X_coordinate	AreaSize	Days	ManHours	Total	TeamSize
2101	6714135.15	1410847.35	87045	185	17001	504	11.49
2103	6694852.90	1425277.18	50772	100	7917	311	9.90
2104	6699519.34	1400965.57	38637	64	8155	339	15.93
2302	6723003.34	1383761.90	180846	215	45205	2340	26.28
2303	6763792.71	1355347.57	135928	202	22938	1202	14.19
2305	6793857.56	1358299.68	99056	64	8629	496	16.85
2602	6711461.67	1459042.77	28724	141	17159	723	15.21
2603	6703954.59	1441172.50	38543	100	10225	377	12.78
2901	6732100.20	1451988.27	71700	186	22999	847	15.46
2904	6738133.40	1440287.36	31252	86	11267	608	16.38
3101	6759393.79	1469808.98	58665	60	8633	238	17.99
3103	6767908.97	1482689.48	53490	74	9433	236	15.93
3104	6798939.34	1473091.66	101254	14	1957	39	17.47
3402	6797349.78	1440870.45	146670	201	18833	479	11.71
3902	6803510.52	1389393.61	259773	224	29862	989	16.66
3904	6839656.56	1350298.61	215877	294	18077	682	7.69
3905	6877628.10	1333752.15	232523	177	9408	351	6.64
6101	6672751.18	1477457.96	37588	34	1949	122	7.17
6102	6653376.12	1481387.44	76360	221	13313	691	7.53
6201	6778189.35	1420440.78	140628	265	27015	1384	12.74
6202	6790296.32	1421891.98	50855	107	7278	255	8.50
6203	6744837.91	1425741.98	47885	54	6320	421	14.63
6205	6754240.67	1396946.38	69850	50	12190	583	30.48
8003	6723790.74	1485616.13	30298	18	1461	63	10.15
8004	6715991.68	1481014.60	3674	29	2917	130	12.57
8005	6711126.55	1499604.51	11459	2	100	5	6.25
8007	6722795.63	1502243.25	18578	30	4465	128	18.60
8008	6737168.61	1511582.04	72562	87	7536	307	10.83
8010	6755556.84	1497572.39	36073	28	2326	83	10.38
8011	6741712.53	1482058.24	19349	5	2024	74	50.60
8102	6701718.34	1472919.76	52095	49	8101	329	20.67
8202	6705146.64	1505368.75	19251	19	1794	104	11.80
8203	6695970.53	1490412.58	8378	7	592	46	10.57
8301	6679997.23	1504565.81	21560	47	5660	202	15.05
8302	6688735.56	1522612.87	23409	42	2699	135	8.03
8303	6703092.31	1516951.19	44049	102	9754	341	11.95
8403	6680642.81	1538305.86	19558	48	2799	131	7.29
8404	6675494.50	1525202.17	23743	34	4648	147	17.09
8406	6676056.93	1519608.57	6930	14	1227	71	10.96
8501	6669554.88	1463179.42	19006	53	5400	225	12.74
8503	6680452.71	1448251.65	71509	115	11892	432	12.93
8504	6668235.64	1430003.75	60214	41	9528	357	29.05

Appendix 2

Neighborhood	Contiguous to:					
2101:	2103	2104	2302	6201	6203	6205
2103:	2101	2603	8504			
2104:	2101	2302				
2302:	2101	2104	2303	3902	6205	
2303:	2302	2305	3904			
2305:	2303	3902	3904			
2602:	2901	8102	8503			
2603:	2103	2901	2904	8503	8504	
2901:	2602	2603	2904	3101	8503	
2904:	2603	2901	6203			
3101:	2901	3103	3104	8011		
3103:	3101	3104	8010	8011		
3104:	3101	3103				
3402:	6201	6202				
3902:	2302	2305	6201	6205		
3904:	2303	2305	3905			
3905:	3904					
6101:	6102	8102	8501			
6102:	6101	8102				
6201:	2101	3402	3902	6202	6203	6205
6202:	3402	6201				
6203:	2101	2904	6201			
6205:	2101	2302	3902	6201		
8003:	8004	8011	8102			
8004:	8003	8102				
8005:	8007	8008	8202			
8007:	8005	8008	8202			
8008:	8005	8007	8010	8202	8303	
8010:	3103	8008				
8011:	3101	3103	8003			
8102:	2602	6101	6102	8003	8004	
8202:	8005	8007	8008	8301	8303	
8203:						
8301:	8202					
8302:	8303	8404	8406			
8303:	8008	8202	8302	8404	8406	
8403:						
8404:	8302	8303	8406			
8406:	8302	8303	8404			
8501:	6101	8503				
8503:	2602	2603	2901	8501	8504	
8504:	2103	2603	8503			