

Sandwich mapping of rodent density in Jilin Province, China

LIU Tiejun^{1,2,3}, *WANG Jinfeng^{1,2}, XU Cheng⁴, *MA Jiaqi⁵, ZHANG Hongyan⁶, XU Chengdong¹

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. China Siwei Surveying & Mapping Technology Co. Ltd., Beijing 100048, China;
4. Prevention and Control Base of Plague and Brucellosis, Chinese Center for Disease Control and Prevention, Jilin 137000, Jilin, China;
5. Chinese Center for Disease Control and Prevention, Beijing 102206, China;
6. School of Geographical Sciences, Northeast Normal University, Changchun 130024, China

Abstract: Rodents are the main host animals that spread plague, and *Spermophilus dauricus* (*S. dauricus*) is the most common rodent in North China. In this study, a rodent density survey was carried out in China's Jilin Province from April to August 2005. Moran's I and semivariogram curves were used to investigate the spatial distribution characteristics of the sampling data. We found that the spatial auto-correlation index was low and failed to generate a meaningful semivariogram curve. In this case, commonly used interpolators, such as kriging, were not suitable for mapping density over the study area. However, the Sandwich model, which is based on spatial stratified heterogeneity, could be applied to our data. Our results showed that the type of soil and land use significantly influenced the distribution of rodent density, and the interactive effect of these variables was much stronger than that of each variable alone. The Sandwich-estimated rodent density map showed that rodent density increased from the southeast to the northwest in Jilin Province. Finally, a framework of a rodent density survey using the Sandwich model was introduced.

Keywords: Sandwich model; geographical detector; *Spermophilus dauricus*

1 Introduction

Plague is a natural epidemic disease with foci that are widely distributed on every continent, excluding Australia (Gage and Kosoy, 2005; Peset, 2015). Given its sudden onset, infectiousness, rapid spread, and high fatality rate, plague is a deadly infectious disease that causes serious damage to humans. Historical records show three incidents of global preva-

Received: 2017-05-05 **Accepted:** 2017-07-27

Foundation: National Natural Science Foundation of China, No.41531179; No.41421001; No.41271404; MOST, No.2016YFC1302504; Special Scientific Research Fund of Public Welfare Profession of China, No.GYHY20140616

Author: Liu Tiejun, PhD, specialized in spatial sampling and data analysis. E-mail: liutj@lreis.ac.cn

***Corresponding author:** Wang Jinfeng, Professor, E-mail: wangjf@lreis.ac.cn

Ma Jiaqi, E-mail: majq@chinacdc.cn

lence of plague, which resulted in at least 160 million deaths (Perry and Fetherston, 1997).

Plague has resurged on a global level since the 1980s. Monitoring results show that the natural epidemic foci and extent of areas affected by plague gradually increase. Plague has been reactivated in animals and still poses a threat to humans. Consequently, in 2000, the World Health Organization listed plague as a resurgent infectious disease. In China, the number of human cases of plague increased from 16 in 1999 to 254 in 2000. It declined after 2000, but the fatality rate multiplied by 25 times overall, reaching 100% in critical areas (Zhou and Yang, 2010). Therefore, the prevention and control of the spread of plague remains an important public health concern. The relationship between the distribution of animals, such as rodents, and the prevalence of plague has long been established (Garnham, 1949). Since various rodents are potential primary hosts facilitating the spread of plague, and rodent density is a common indicator in evaluating the risk of a plague outbreak (Lotfy, 2015; Wang and Li, 1997; Yang *et al.*, 2000), monitoring and the statistical inference of this indicator directly affect the evaluation of the risk of plague outbreaks (Gage and Kosoy, 2005; Stenseth *et al.*, 2008).

Several scholars have studied rodent distribution (Jiang *et al.*, 2007; Liu *et al.* 2007; Li *et al.*, 2012; Zhao, 2012; Wang *et al.*, 2013). In these studies, the impact of geographic changes (e.g., soil, vegetation, and terrain) on rodent density was considered when determining monitoring spots; however, most of them focused on the relationship between rodent density and its environmental factors, few interpolated maps are presented, probably due to the high spatial stratified heterogeneity of the target, which is different from the principle of spatial autocorrelation assumed by most existing interpolation models. Fan *et al.* (2016) used the empirical Bayesian kriging model to map the spatial distribution of *Spermophilus dauricus* (*S. dauricus*) in Manchuria, but they did not evaluate the suitability of such a model in their study—after all, spatial statistics, such as kriging and geographically weighted regression (GWR), are believed to only be suitable for the modeling of inherently continuous features (Diggle, 2003). Besides geostatistical techniques, point pattern analysis (such as kernel smoothing techniques) (Wand and Jones, 1994), generalized linear models (GLM) (Thuiller, 2003), and ecological-niche factor analysis (ENFA) (Hirzel and Guisan, 2002), etc., are also widely used. Moreover, Hengl *et al.* (2009) suggested an optimized method by combining the point pattern analysis, ENFA, and GLM techniques, which was believed to be more promising as it would exploit the advantages of each. However, its output map was found to be too greatly affected by the distribution of the samples rather than the environmental predictors, while the predictors could be accurately selected, the result would likely be biased if the samples' distributions were unjustified. Furthermore, if rodent density were distributed heterogeneously, and the values of the environmental predictors were categorical variables (such as types of vegetation or land use), then it would be difficult to implement the GLM process if there were too many categories (Salim and Daoud, 2003).

The Sandwich model (Wang *et al.*, 2002, 2013) is a simple and applied technique for the study of subjects with remarkably distinct spatial characteristics, and has been used in the fields of environmental survey and public health etc. Chen *et al.* (2012) analyzed soil *Cr* content according to the concept of the Sandwich model at the stage of setting up samples, and they found that the samples obtained by Sandwich sampling achieved the minimum kriging estimation error. Wu *et al.* (2016) performed a regional precipitation survey in the same way as Chen *et al.* (2012), where Sandwich sampling was also found to be the best

approach. However, at the statistical interpolation stage, both sets of researchers relied on kriging to provide an interpolation model rather than the Sandwich model; this might be caused by the researchers' inexperience with the Sandwich model. Hu *et al.* (2015) mapped the schistosomiasis risk in China's Anhui Province by using the Sandwich model with a knowledge layer of land use, which was a well-recognized factor closely related to the distribution of schistosomiasis, whereas the geographical detector model proposed by Wang *et al.* (2010b) was used to check the spatial stratification of the schistosomiasis risk. The results showed that the spatial distribution of land use could explain 71% of the risk variation.

Northwest Jilin Province, China, is the natural focus of plague, and *S. dauricus* is its main source of plague and a harmful rodent species in this location (Zhang *et al.*, 2004; Zhang *et al.*, 2006; Jiang *et al.*, 2007; Liu *et al.*, 2007). In the present study, taking four municipal cities in western Jilin as study regions, a sample survey was performed to investigate the distribution of *S. dauricus* density in the study regions by using the Sandwich model. Similar to Hu *et al.*'s (2015) study, the geographical detector model (Wang *et al.*, 2010b) was used to select the best influencing factors of rodent distribution which were then taken as knowledge layers before implementing the Sandwich model. However, this study went one step further to optimize the knowledge layer by studying the interaction effects of alternative factors, according to the concept of geographical detectors. Additionally, a framework was developed to streamline the general process of carrying out a rodent density survey through the use of the Sandwich model, which could provide guidance to future researchers.

This paper is organized as follows: Section 2 describes the characteristics of the rodent survey data and the methods used in estimating multiregional rodent density and obtaining optimized knowledge layers of the Sandwich model. Section 3 presents the Sandwich mapping results. Section 4 presents the discussion and concluding remarks.

2 Data and method

2.1 Rodent survey data

In this study, 50 survey areas were selected randomly in the study area. The study was conducted from 2nd April to 7th August, 2005. The quadrat method was used to survey rodent density in each of the survey areas. Within each survey area, 0.5% of the sampling quadrats were 100 m * 100 m in size and were selected by considering their differences in geomorphology. In each quadrat, one bow clip (see Figure 1a) was installed at each rodent hole and was checked every 2 hours from sunrise to sunset every day. The bow clips with caught rodents were replaced with new ones. Thus, rodent density was defined as the number of rodents caught in each quadrat in one day.

Rodent populations in arid and semi-arid regions usually vary according to the season (Liu *et al.*, 2007). *S. dauricus* is a dormant species that comes out of hibernation in April and enters hibernation

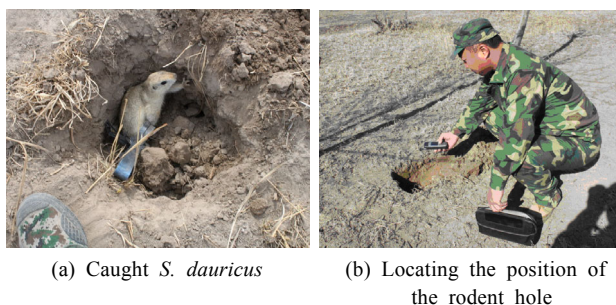


Figure 1 Photographs of the field survey

in early September. The species reaches its yearly population minima in early April and achieves it maxima between July and August (Luo and Zhong, 1990). To avoid the influence of seasonal change on the modeling result, survey data from July with 627 sampling quadrats distributed in 14 survey areas were chosen from the total of 2676 samples to compute the entire study area’s rodent density.

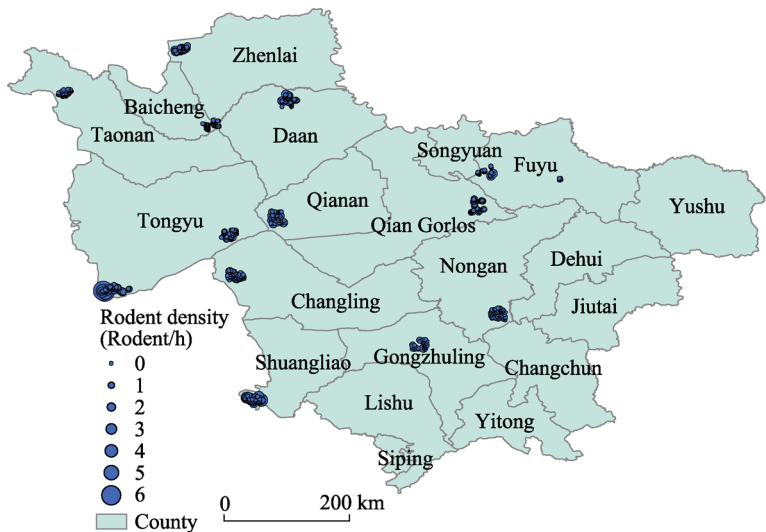


Figure 2 Distribution of rodent density from survey data

Figure 2 illustrates the positions of the survey points. Rodent density at a sample point is represented by the size of that point. Samples with the highest rodent density were found in Tongyu. Figure 3 is a histogram of sampled rodent density; more than 92% samples are zero, and only two samples are 6.

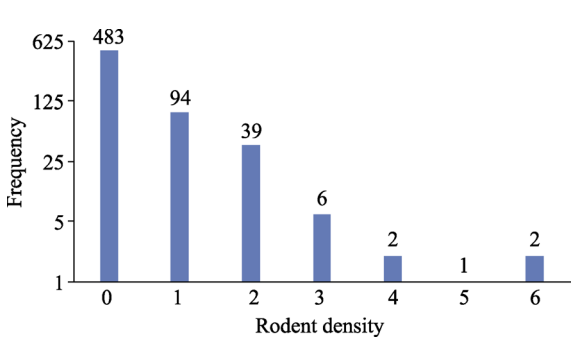


Figure 3 Histogram of rodent density samples

2.2 Analysis of the spatial distribution characteristics

Testing the distribution of the spatial characteristics of the study subjects is essential for selecting an appropriate sampling and statistical method. First, the Moran’s I of the samples was investigated to observe the autocorrelation of rodent density distribution. The spatial adjacency matrix used in the calculation was established using the inverse distance weighting method, that is, a farther distance corresponds to a smaller weight. In the present study, the Moran’s I value was 0.24, which means that the spatial correlation of the rodent density survey data was not high.

Furthermore, the semivariogram curve of the survey data was obtained through a spherical model (Figure 4) to determine the distribution pattern of the samples. The variation function curve showed that no stable semivariogram could be established to model the distribution of the rodent density samples.

The results of the exploratory data analyses conducted indicated that kriging, which relies

on global spatial correlation, was not suitable for analyzing our data. This might be a consequence of the conditions under which the sampling quadrats were aggregated. In the 14 scattered areas sampled (see Figure 2), the average distance between each pair of the nearest neighbored quadrats was about 2 km in each area, whereas the distance between the nearest neighbored areas was 30 km; thus, the semivariance curve could not be modeled because of a lack of data in the medium range. Thus in this analysis, we used the Sandwich model which relies on spatial heterogeneity.

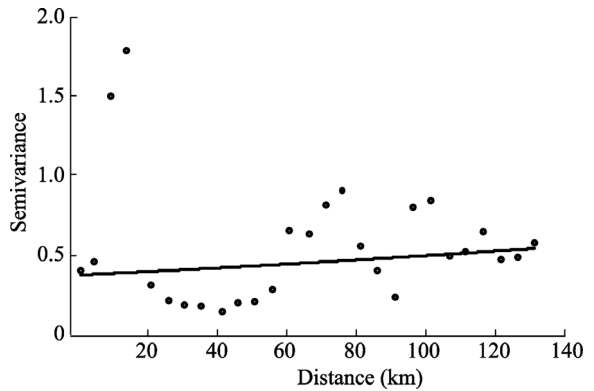


Figure 4 Semivariogram curve of rodent density survey data in Jilin (Spherical model, Nugget 0.30, Partial Sill 0.69, Range 468.185.40)

2.3 Geographical detector

The geographical detector model can be used to quantitatively assess spatial heterogeneity and the influence of potential driver factors (Wang *et al.*, 2010b; Hu *et al.*, 2015). For instance, in this study it was used to detect the influence (qualified by the q -value) of a set of selected knowledge layers on the distribution of rodent density. The higher the q -value, the stronger the determinant power of the knowledge layer. The q -value is found to follow the noncentral F probability density distribution (Wang *et al.*, 2016). Besides the factor detector, the interactive detector is given by the geographical detector, which is used to evaluate the combined determinant power of the two knowledge layers.

The q -value of a given knowledge layer can be expressed as follows:

$$q = 1 - \frac{1}{N\sigma^2} \sum_{h=1}^L N_h \sigma_h^2 \quad (1)$$

where σ^2 refers to the dispersion variance in the entire study region, N_h refers to the number of subpopulations in the knowledge unit, N refers to the number of populations in the whole region. σ_h^2 is the dispersion variance in knowledge unit h . The second item at the right side of the equation is the weighted sum of variance of each knowledge unit. $q \in [0, 1]$; if the attributes of all subpopulations in each of the knowledge units are the same, then q is equal to 1, if the population is random, then the second item at the right side of the equation is 1, and q is equal to 0. With respect to the interactive detector, a new knowledge layer is generated by overlaying two given knowledge layers; thus, h and L in formula (1) are the unit and unit number of the new knowledge layer, respectively.

2.4 Sandwich model-based mapping

The Sandwich mapping model is based on the concept of spatial stratified heterogeneity. It consists of sampling, knowledge, and reporting layers. The knowledge layer is essential to determine the accuracy of the statistical inference for the Sandwich model, and this layer can

be established using a priori knowledge, historical data, or proxy variable analysis (Li *et al.*, 2008; Wang *et al.*, 1997). In this study, this model established the transfer of information and variance from the sampling quadrats to the reporting layer (county regions) via knowledge layers to eliminate the stiff link between sampling and statistical inference.

First, the mean value of rodent density \bar{y}_z and its variance $v(\bar{y}_z)$ of the z^{th} unit of the knowledge layer were calculated. Then, the mean value and its variance of the r^{th} county were calculated from the knowledge layer on the basis of the feature of stratified sampling (Cochran, 1977),

$$\bar{y}_r = \frac{1}{N_r} \sum_{z=1}^{L_r} N_{rz} \bar{y}_z = \sum_{z=1}^{L_r} W_{rz} \bar{y}_z \quad (2)$$

where $r = 1, \dots, N_r$, and W_{rz} refers to the weight of the z^{th} unit of knowledge layer in the r reporting unit.

The total observed variance of reporting unit r can be expressed as:

$$v(\bar{y}_r) = \sum_{z=1}^{L_r} W_{rz}^2 v(\bar{y}_z) \quad (3)$$

Hence, the mean and its variance in the reporting layer were estimated on the basis of the knowledge layer instead of the sampling layer.

2.5 Selection and optimization of knowledge layers

As discussed earlier, the performance of Sandwich mapping is determined by the knowledge layers that are generated by prior knowledge of the spatial distribution of *S. dauricus*. Generally, *S. dauricus* thrives in habitats with broad, non-cultivated, and overgrown dunes, in dwarf crop lands on either side of earthen roads, in field spaces, and in meadow steppes. The density of *S. dauricus* varies depending on the soil, vegetation, and geomorphology. Recently, human activity was found to dramatically affect the density of *S. dauricus* (Gang *et al.*, 2006; Zhou *et al.*, 2007; Luo and Zhong, 1990). Rodent density is high where soil texture and terrain are appropriate, and where plants and fruits serve as rich food sources. Brown calcic soil in gently sloping *Caragana microphylla* hills, and areas with tuft grass and *Artemisia frigida* are conducive habitats for *S. dauricus*. Additionally, soil water conditions were found to have a negative relationship to *S. dauricus* density, as the species prefers short grass or dry soil habitats (Wang *et al.*, 2003). Among the abovementioned factors, the datasets of soil, land use, vegetation, and Normalized difference vegetation index (NDVI) were selected as alternative knowledge layers (Figure 5), where soil data reflects the distribution of different soil types, land use data represents conditions of geomorphology and human activity (since land use is the outcome of the interaction between geomorphic conditions and human activity, etc.), vegetation data reflects the distribution of different vegetation types, and NDVI data reflects the abundance of vegetation cover.

The q-values and their statistical significances of the four knowledge layers and their interactive effects are shown in Table 1.

It is noteworthy that there will be too many units generated in new knowledge layers after overlaying two given layers, and that a lot of units will have no samples, thus the q-value will fail to be calculated. To avoid such results, a two-step cluster command was performed in SPSS to merge the knowledge units. The unit number of combined knowledge layers was

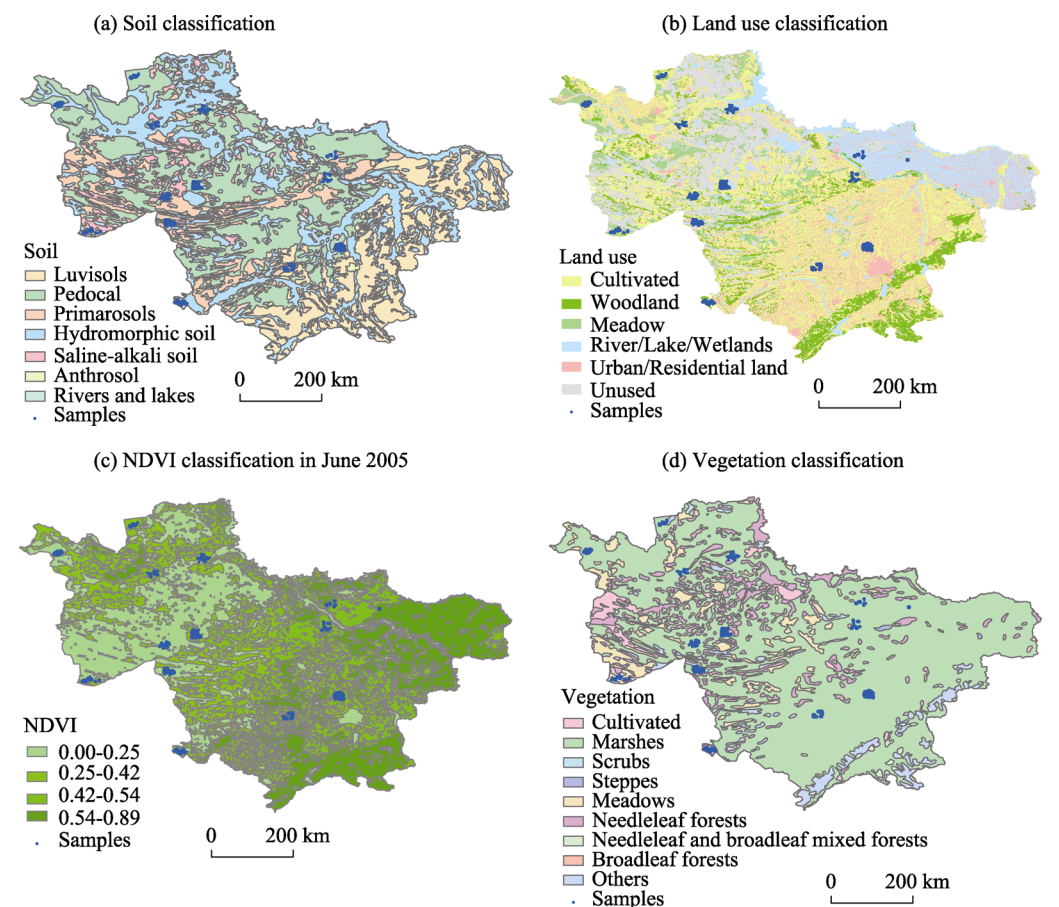


Figure 5 Alternative knowledge layers in the study area

set to nine, which was determined by the unit number of the vegetation classification with the most units among the four knowledge layers. The significance of the q-value was calculated following Wang *et al.* (2016).

A comparison of the q-values of the four knowledge layers implies that soil(0.323) and land use(0.288) classifications are the top two; the summaries of samples laid in each unit of the two layers are listed in Tables 2 and 3. Intuitively, soil is the preferred knowledge layer in this study, and land use is secondary. However, the combined q-value of soil and land use (0.412) is much higher than that of soil alone, which

Table 1 q-values of single variables and their interactive effect

	Soil	Land use	NDVI	Vegetation
Soil	0.323***			
Land use	0.412***	0.288***		
NDVI	0.237***	0.347***	0.255***	
Vegetation	0.374***	0.296***	0.290***	0.168***

Significance level: *** 0.01

Table 2 Summary of samples laid in the soil layer

Name	Sample count	Mean (rodent/h)	Variance
Luvisols	14	0.07	0.07
Pedocal	297	0.29	0.34
Primarosols	153	0.48	0.7
Hydromorphic	118	0.12	0.19
Saline-alkali	41	0.98	2.52
Anthrosol*	0	0	0
Lake/River	4	0	0

*Since the anthrosol area accounted for only 1.2% of the study area, there were no anthrosol samples which were neglected in the q-value and Sandwich calculations.

means that the determinant power is raised significantly by the combined effect of the two layers. Consequently, it is reasonable to optimize the knowledge layer of the Sandwich estimation of the rodent density by combining the soil and land use layers (Figure 6).

Table 3 Summary of samples laid in the land use layer

Name	Sample count	Mean (rodent/h)	Variance
Cultivated	305	0.16	0.21
Woodland	61	0.77	0.95
Meadow	134	0.59	0.88
River/Lake/Wetlands	51	0.1	0.17
Urban area	19	0	0
Unused	57	0.61	1.31

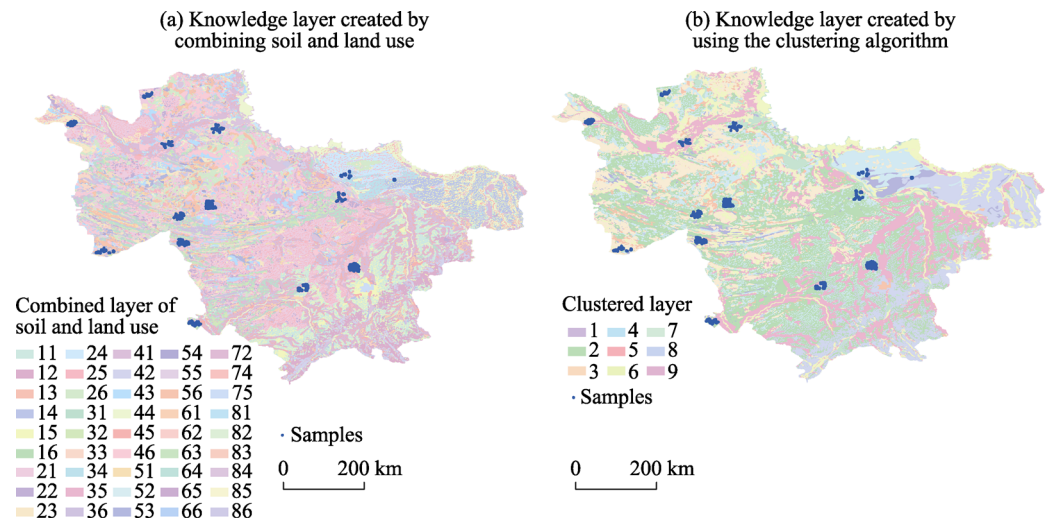


Figure 6 Combination of soil and land use layers

3 Sandwich mapping

The results of the Sandwich mapping of rodent density in 20 county administrative units in the study area, which were generated by Sandwich software found on the website www.sssampling.org/sandwich, are shown in Figures 7–10, where four types of knowledge layers, including soil, land use, NDVI, and vegetation type, were used. The domain division in the mean statistical thematic maps considered the comprehensive minimum (0.08) and maximum (0.5) values as the initial value and final value, respectively. Then, eight division levels were set at the average interval. In this way, the colors of all mean thematic maps were comparable, and the differences in the spatial distributions of the estimated results between the different knowledge layers were expressed.

Typically, the reliability of an estimate ($\hat{\mu}$) is assessed by mean squared error (MSE) (Cochran, 1977: 15), which is defined as

$$MSE(\hat{\mu}) = V(\hat{\mu}) + (B)^2 \tag{4}$$

where $V(\hat{\mu})$ is the variance of $\hat{\mu}$, and B is the sampling bias. Since the Sandwich model is based on stratified sampling (Wang *et al.*, 2013) that is theoretically unbiased (Cochran, 1977: 91), i.e. $B = 0$, the variances of estimates $V(\hat{\mu})$ generated by the four knowledge layers of the Sandwich model could be used as the indicators of their performance, where a smaller variance represents a more reliable estimate (Cochran, 1977; Haining, 1988; Griffith *et al.*, 1994; Wang *et al.*, 2010a; Wang *et al.*, 2013). In this study, the variances of the statis-

tical thematic maps were presented together with the mean map. The variance maps were drawn in the same way as the mean map, with the minimum value being 0.04 and maximum value being 0.5.

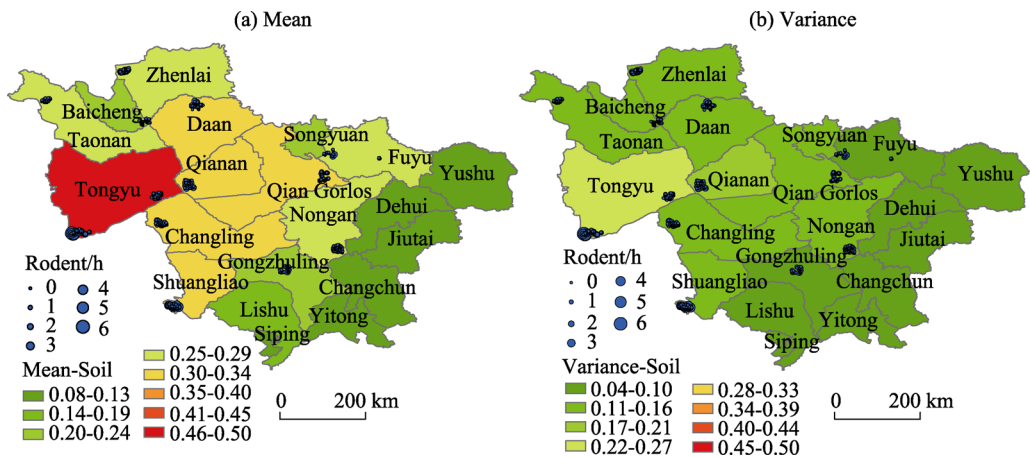


Figure 7 Sandwich estimating result based on the soil knowledge layer

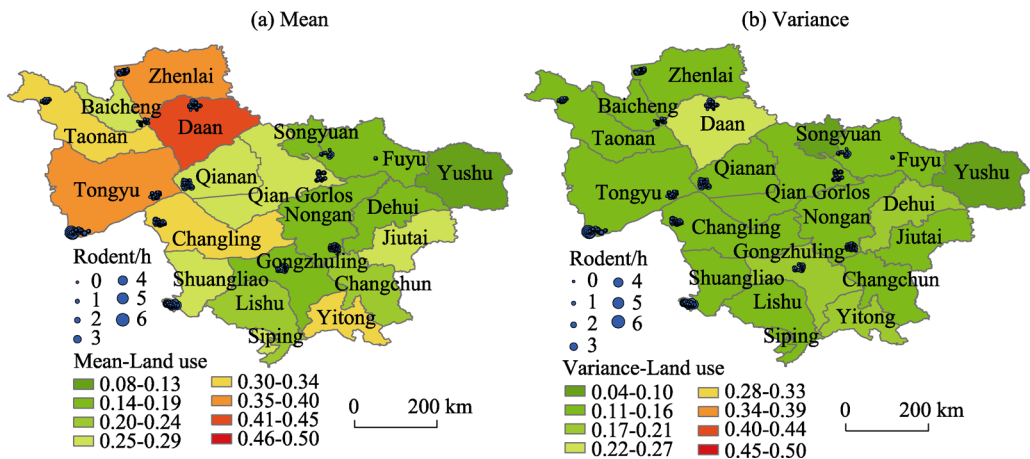


Figure 8 Sandwich model estimates based on the land use knowledge layer

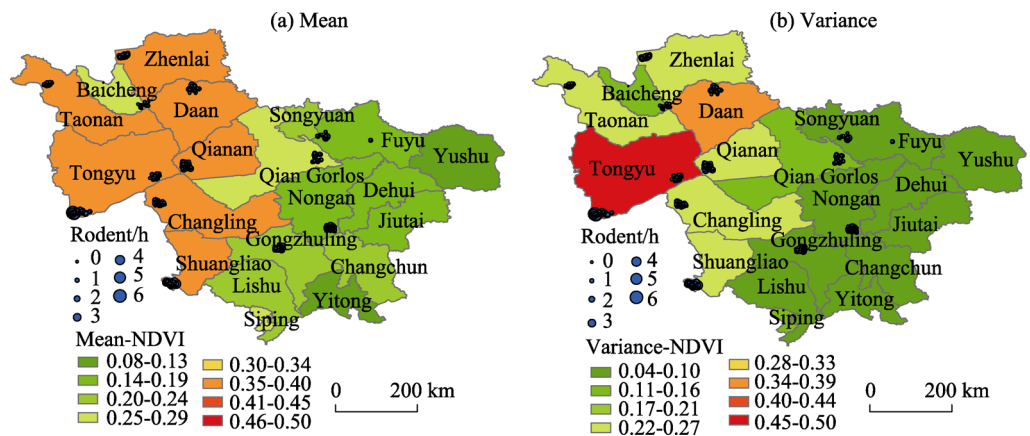


Figure 9 Sandwich model estimate result based on the NDVI knowledge layer

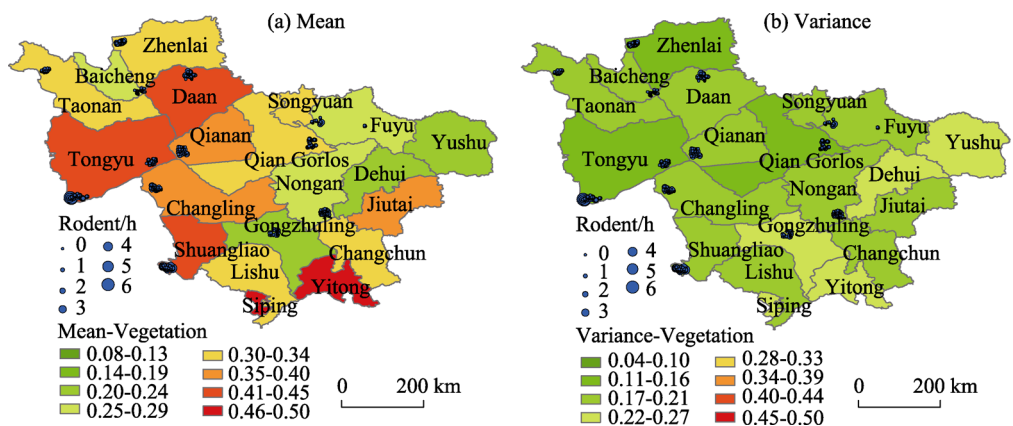


Figure 10 Sandwich model estimate based on the vegetation knowledge layer

Comparative analysis of the statistical results from the four knowledge layers showed that the vegetation layer yielded, on average, the highest variance (average variance of the mean of all the counties reached 0.192), followed by the NDVI (0.149), and the land use (0.134) layers. The soil type layer yielded the lowest average variance (0.099). Generally, the reliability of the estimated values from the four knowledge layers were in the order of soil type > land use > NDVI > vegetation type, which was the same as the order of their q-values. This finding highlighted the fact that the q-value could be used as a reference in finding suitable knowledge layers during the process of Sandwich mapping.

As mentioned before, combining the soil type and land use layers might optimize the Sandwich mapping of rodent density. The thematic maps of the estimated means and variances from the combined knowledge layers are shown in Figure 11. The classification rule is consistent with that of the four previous knowledge layers.

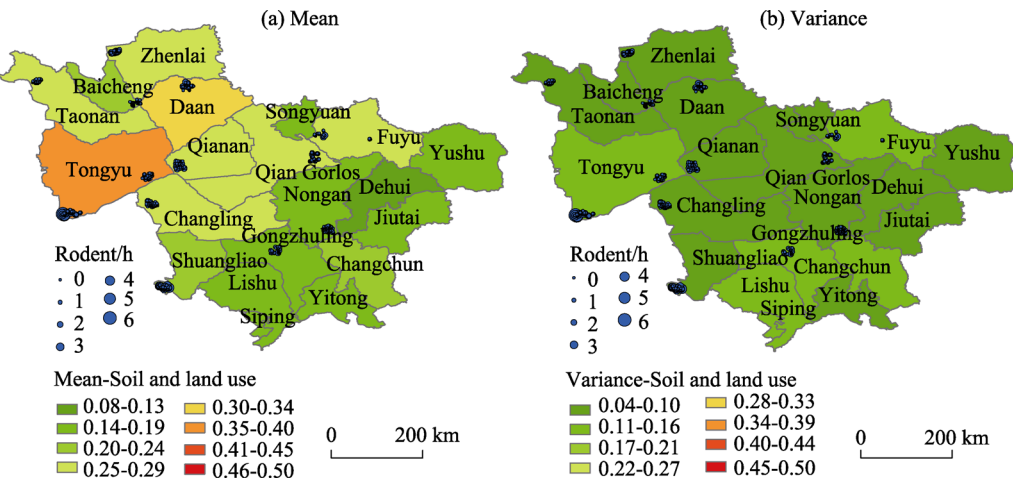


Figure 11 Sandwich model estimates based on the combined knowledge layer

A comparison of the statistical variance map between the Sandwich model estimates from the combined knowledge layers and the four single layers shows that the variance of the estimated mean values in the whole study area from the combined knowledge layer (0.085) is less than that of the individual layers. This indicates that the results from the combined

knowledge layers are the most reliable.

From the mean statistical map using the combined knowledge layer, the overall trend in the distribution of rodent density was higher in the northwestern counties than it was in the southeastern counties. The rodent densities in each of the counties varied from 0.122/h to 0.373/h, with an area weighted mean of 0.22/h. Tongyu, Daan, Changling, Zhenlai, and Taonan counties, located in the northwest of the study area, were the top five counties with the highest rodent density. Rodent densities of counties without samples, such as Lishu, Siping, Jiutai, etc., were still estimated using the Sandwich model based on the bridge of the knowledge layers. This result was similar to that of both Jiang (2007) and Zhou (2007) who studied rodent density survey data in 16 counties in northwestern Jilin from 2005 to 2006 and found that the counties of Tongyu, Qianan, Taonan, and Shuangliang (located in the northwest of the study area) had the highest rodent density; the mean rodent density was 0.27/h. It is noteworthy that the results differed a little from ours, probably because the survey periods were different; additionally, the mean value in our study was calculated indirectly from the knowledge layer of the Sandwich model, instead of from the samples themselves.

4 Discussion and conclusions

This study introduced an applied method of mapping rodent density in 20 counties in northwestern Jilin by considering the ecological habits and distribution law of *S. dauricus*—the main rodent species in the entire study area. Soil type was found to be the most reliable explanatory variable for Sandwich model estimates of rodent density; however, the performance of the combined layers of soil type and land use was even better than that of soil type alone. The advantages of this study in comparison with previous studies are discussed below.

First, since the spatial distribution of the sampling quadrats was not even, spatial correlation was impractical for the estimation of rodent density in blank areas; hence, spatial correlation-based models, such as kriging, were not suitable for this study (Wang *et al.*, 2012; Wang *et al.*, 2013). However, the distribution pattern of samples was not a problem for the Sandwich model. Unlike kriging, the Sandwich model can incorporate spatial heterogeneity by following a concept similar to stratified sampling, which assumes that the inner value of a given stratum (unit of knowledge layer) is homogeneous (Cochran, 1977). In this way, sampling information can be transferred from areas with samples to areas without samples via the knowledge layer; as a result, the rodent density of counties (units of reporting layer) without samples laid can also be reliably inferred.

Secondly, the geographical detector model was used in this study to evaluate the influence of various natural and human factors on rodent distribution and their interactive effects. By comparing *q*-values, the factors with the greatest effects were selected (Huang *et al.*, 2014), which was important for setting sampling quadrats and for the Sandwich model estimation. By extrapolating these results, the next time when a new rodent density survey is performed in an area like northwestern Jilin Province, given the existing knowledge layers of soil type and land use, only a few sampling quadrats will be needed in each unit of the knowledge layer, because the rodent density across one unit is thought to be approximately the same. In the Sandwich model estimation, the variances of the estimated means of the reporting units

were determined by the quality of the knowledge layer (Wang *et al.*, 2010a). The geographical detector model is an applied tool to ensure the accuracy of the Sandwich model.

Finally, the Sandwich model used in this study fully considered the priori knowledge of rodent habitat and community structure. We started with prior knowledge of the natural and human factors that were closely related to the distribution of rodents. Then, we used the sampling data to verify the representativeness of the knowledge layer instead of the sample itself. This process effectively used previous knowledge and incorporated it into the calculation process of the model, thus avoiding bias and reducing our dependence on data sampling (Li and Wang, 2004; Zhang *et al.*, 2013). Therefore, the reliability of the obtained result was high.

The most significant result of this study was that we used Sandwich mapping to extrapolate our sample data to 20 counties, and every county was assigned a rodent density of its own. Thus, local departments of disease control could distribute their deratization material according to the differences in rodent density in each county. This is an efficient means of increasing the effectiveness of deratization, and to prevent the unnecessary wastage of materials.

However, there are also some limitations to this study. Although the results of this study indicate that soil and land use and their combined effect are the most appropriate knowledge layers for *S. dauricus* surveying, there were several factors worth further study. For example, we compared the q-values of some other factors, such as geomorphology and terrain, which were also thought to be closely related to the distribution of *S. dauricus* in the study area, but their q-values were much lower than the four factors mentioned above. It may be because the distribution of the quadrats did not fit the zoning of the geomorphology and terrain factors well, thus the inner variances of the units of the factors were much too great for q-value calculation.

In summary, in the fields of public health and ecological environment, etc., several concerns have been raised with regard to applying sampling and inferring models, including how to select the most suitable models, and how to obtain the best results under the existing conditions. Particularly for users who are unfamiliar with sampling methods and theories,

these problems impede scientific and stable applications of spatial sampling methods. Based on the findings of this study, a formal process of engaging in a rodent density survey could be raised to answer the questions above, as shown in Figure 12.

In undertaking a task such as a rodent density survey, the first step is to collect prior knowledge, including the spatial distribution law of the rodent and the factors closely related to it. If there are ex-

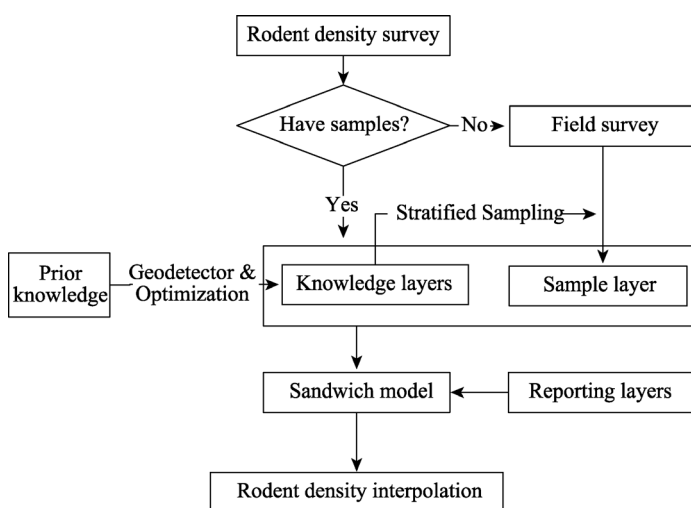


Figure 12 The rodent density survey framework using the Sandwich model.

isting samples, then the second step is to utilize Geodetector software to select the best knowledge layer, or optimize it by combining the layers with the largest interactive q-values. Otherwise, stratified sampling should be performed according to the knowledge layer that was chosen based on the prior knowledge. Finally, the sampling, knowledge, and reporting layers are prepared for the Sandwich model, and the mapping of rodent density is obtained.

References

- Chen F, Dai H, Hu Y *et al.*, 2012. Study on regional soil spatial sampling method. *Geography and Geo-Information Science*, 28(6): 53–56. (in Chinese)
- Cochran W G, 1977. Sampling Techniques. 3rd ed. New York: Wiley.
- Diggle P, 2003. Statistical Analysis of Spatial Point Patterns. 2nd ed. Arnold Publishers.
- Fan L, Wu E, Qu X *et al.*, 2016. Distribution characteristics of *Spermophilus dauricus* in Manchuria City in China in 2015 through ‘3S’ technology. *Biomed Environ. Sci.*, 29(8): 603–608.
- Gage K, Kosoy M, 2005. Natural history of plague: Perspectives from more than a century of research. *Annual Review of Entomology*, 50(1): 505–528.
- Gang S, Liu Z, Zhang Y *et al.*, 2006. A study on the landscape characteristics of plague infectious areas in Jilin Province. *Chinese Journal of Control of Endemic Diseases*, 21(5): 257–259. (in Chinese)
- Gang S, Liu Z, Zhou F *et al.*, 2007. Study on present conditions of Jilin province plague nature focus. *Chinese Journal of Control of Endemic Diseases*, 22(3): 161–166. (in Chinese)
- Garnham P C, 1949. Distribution of wild-rodent plague. *Bulletin of the World Health Organization*, 2(2): 271.
- Griffith D A, Haining R, Arbia G, 1994. Heterogeneity of attribute sampling error in spatial data sets. *Geographical Analysis*, 26: 301–320.
- Haining R, 1988. Estimating spatial means with an application to remote sensing data. *Communication Statistics: Theory and Methodology*, 17: 537–597.
- Hengl T, Sierdsema H, Radović A *et al.*, 2009. Spatial prediction of species’ distributions from occurrence-only records: Combining point pattern analysis, ENFA and regression-kriging. *Ecological Modelling*, 220(24): 3499–3511.
- Hirzel A, Guisan A, 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, 157 (2–3), 331–341.
- Huang J, Wang J, Bo Y *et al.*, 2014 Identification of health risks of hand, foot and mouth disease in China using the geographical detector technique. *International Journal of Environmental Research and Public Health*, 11(3): 3407–3423.
- Hu Y, Bergquist R, Lynn H *et al.*, 2015. Sandwich mapping of schistosomiasis risk in Anhui Province, China. *Geospatial Health*, 10(1): 324.
- Li B, Zhang M, Wang Y *et al.* 2012. Surveys on the population densities of rodent communities in some hilly areas of Changning of Hunan and Xinyu of Jiangxi in China in 2010–2011. *Plant Protection*, 38(1): 155–157. (in Chinese)
- Li L, Wang J, 2004. Integrated spatial sampling modeling of geospatial data. *Science China Earth Sciences*, 47(3): 201–208.
- Li L, Wang J, Cao Z *et al.*, 2008. An information fusion method to regionalize spatial heterogeneity for improving the accuracy of spatial sampling estimation. *Stochastic Environmental Research and Risk Assessment*, 22: 689–704.
- Liu W, Wang X, Zhong W, 2007. Population dynamics of the Mongolian gerbils: Seasonal patterns and interactions among density, reproduction and climate. *Journal of Arid Environments*, (68): 383–397.
- Liu Z, Zhang Y, Zhang F *et al.*, 2007. Study on distribution scope and accumulation areas of *Spemophilus dauricus* on various administrative division in the west of Jilin Province. *Chinese Journal of Control of Endemic Diseases*, 22(1): 11–13. (in Chinese)
- Lotfy W, 2015. Plague in Egypt: Disease biology, history and contemporary analysis: A minireview. *Journal of*

Advanced Research, 6(4): 549–554.

- Luo M, Zhong W, 1990. Some data of *Citellus dauricus* population ecology. *J. Anim. Sci.*, (2): 50–54. (in Chinese)
- Perry R D, Fetherston J D, 1997. *Yersinia pestis*-etiologic agent of plague. *Clin. Microbiol. Rev.*, 10(1): 35–66.
- Peset J, 2015. Plagues and Diseases in History. International Encyclopedia of the Social & Behavioral Sciences. 2nd ed. 2015: 174–179.
- Salim S, Daoud J I, 2003. Regression Analysis with Dummy Variables. Understanding Regression Analysis. New York: Springer US.
- Stenseth N, Atshabar B, Begon M *et al.*, 2008. Plague: Past, Present, and Future. *PLoS Medicine*, 5(1): e3.
- Thuiller W, 2003. BIOMOD: Optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9(10): 1353–1362.
- Wand M P, Jones M C, 1994. Kernel smoothing. *Biometrics*, 54.
- Wang G, Zhong W, Zhou Q *et al.*, 2003. Soil water condition and small mammal spatial distribution in Inner Mongolian steppes, China. *Journal of Arid Environments*, (54): 729–737.
- Wang J, Jiang C, Hu M *et al.*, 2013. Design-based spatial sampling: Theory and implementation. *Environmental Modelling & Software*, 40(40): 280–288.
- Wang J, Haining R, Cao Z, 2010. Sample surveying to estimate the mean of a heterogeneous surface: reducing the error variance through zoning. *International Journal of Geographical Information Science*, 24(4): 523–543.
- Wang J, Haining R, Liu T *et al.*, 2013. Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning A*, 45(10): 2515–2534.
- Wang J, Liu J, Zhuang D *et al.*, 2002. Spatial sampling design for monitoring cultivated land. *International Journal of Remote Sensing*, 23: 263–284.
- Wang J, Li X, Christakos G *et al.*, 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*, 24(1): 107–127.
- Wang J, Stein A, Gao B *et al.*, 2012. A review of spatial sampling. *Spatial Statistics*, 2: 1–14.
- Wang J, Wise S, Haining R, 1997. An integrated regionalization of earthquake, flood and drought hazards in China. *Transactions in GIS*, (2): 25–44.
- Wang J, Zhang T, Fu B, 2016. A measure of spatial stratified heterogeneity. *Ecological Indicators*, 67: 250–256.
- Wang W, Jiang H, Zhao Y *et al.*, 2013. Analysis on the rat density monitoring results, Qingdao City, 2006–2012. *Preventive Medicine Tribune*, 19(10): 758–759. (in Chinese)
- Wang W, Li R, 1997. The main fields of study on environment health and development in 21st century in China. *Progress in Geography*, 16(1): 11–14. (in Chinese)
- Wu J, Liu S, Ma L, 2016. Comparison analysis of sampling methods to estimate regional precipitation based on the kriging interpolation methods: A case of northwestern China. *Sciences in Cold and Arid Regions*, 8(6): 485–494.
- Yang L, Chen R, Wang W *et al.*, 2000. The temporal and spatial distribution of the plague foci since 1840 in China. *Geographical Research*, 19(3): 243–248. (in Chinese)
- Zhang C, Lv J, Pu Q, 2004. The present situation and control countermeasure of *Spermophilus Dauricus* plague natural foci. *Chinese Journal of Control of Endemic Disease*, 19(6): 345–348. (in Chinese)
- Zhang G, Zhang G, Zhang F *et al.*, 2006. A study on the space structure of plague natural infectious areas of Jilin Province. *Chinese Journal of Control of Endemic Disease*, 21(4): 193–196. (in Chinese)
- Zhang J, Zhang L, Gong B, 2013. Application of spatial sampling to remote sensing monitoring of forest cover area. *Advanced Materials Research*, 610–613: 3732–3737.
- Zhao A, 2012. Analysis on monitoring result of rat density in Taian in 2010. *Chinese Journal of Hygienic Insecticides & Equipment*, 18(1): 28–29. (in Chinese)
- Zhou F, Liu Z, Zhang G *et al.*, 2007. Study on *Spemophilus Dauricus* gathering areas of plague natural focus in Jilin Province. *Modern Preventive Medicine*, 34(19): 3648–3652. (in Chinese)
- Zhou D, Yang R, 2010. Progress and prospect of research work on plague. *Medical Journal of Chinese Peoples Liberation Army*, 35(10): 1176–1182. (in Chinese)