Spatial and Temporal Epidemiologic features analysis of pulmonary tuberculosis in Nanjing, China

Hongyu Zhao
Department of Clinical Research Center, The Second Hospital of Nanjing, Nanjing University of Chinese Medicine
Changkui Sun
Department of Remote Sensing Imagery, Provincial Geomatics Center of Jiangsu
Wei Ma
Department of Clinical Research Center, The Second Hospital of Nanjing, Nanjing University of Chinese Medicine
Guoping Yin
Department of Clinical Research Center, The Second Hospital of Nanjing, Nanjing University of Chinese Medicine
Xia Zhang
Department of Clinical Research Center, The Second Hospital of Nanjing, Nanjing University of Chinese Medicine
Huiyan Yu (yuhy86@126.com)
Jiangsu Province Center for Disease Control and Prevention https://orcid.org/0000-0001-9507-8414
Jinyun Song
Department of Clinical Research Center, The Second Hospital of Nanjing, Nanjing University of Chinese Medicine

Research article

Keywords: Pulmonary tuberculosis, Spatial epidemiology, Maxent, Environmental factors, Descriptive statistics

Posted Date: December 18th, 2019

DOI: https://doi.org/10.21203/rs.2.19145/v1

Abstract

Background This study aims to analyse the epidemiological features of tuberculosis in Nanjing and to identify possible risk factors associated with the spatial-temporal distribution.

Methods Firstly, we used descriptive statistical method to study the epidemiologic features of confirmed PTB cases in 2017. Secondly, we explored Maxent method to construct the species distribution model with a high spatial resolution of 1km based on those confirmed cases and environmental features. Once again, we validated the performance of the species distribution model using ROC approach and spatial jackknife test, identified the key environmental factors associated with PTB occurrence. Finally, we created a prediction map under specific environmental factors by projecting the training model onto the research areas.

Results The seasonal amount of PTB was not obvious. The key risk factors associated with PTB occurrence are monthly mean vapor pressure, solar radiation, monthly average temperature, and precipitation with suitable ranges. The economically developed central city is the high-burden area of tuberculosis with a high risk probability, and the economically backward Gaochun district is the high transmission area of tuberculosis with a high incidence. Epidemiologically, the farmers, the elderly, the floating population, and males are the focus groups for PTB prevention and control.

Conclusion The combination of environmental factors with Maxent methods is an appropriate option to analyse and estimate the spatial and temporal distribution of PTB cases.

Background

Pulmonary tuberculosis (PTB) is a chronic infectious disease caused by Mycobacterium tuberculosis, and remains one of the major causes of death among infectious diseases. China is one of the countries with a high burden of PTB in the world [1]. At present, China's PTB prevention and control effect is significant. The overall incidence is on the decline, but the annual number of reported cases is still large, especially in remote and poor areas [2].

Since the 1990s, spatial analysis technology has been widely used in disease mapping, spatial and temporal surveillance of infectious diseases, environmental health, distribution and optimization of medical service resources [3, 4]. Among those spatial geographic analysis models, ecological niche models (ENM) can be used to explore the non-random relationship between disease and environmental factors, to study the spread rule of infectious diseases based on the case information of known vectors, hosts, pathogens, populations and relevant environmental data [5]. Those methods have been utilized by researchers as a useful tool to characterize the potential ecological distribution of species, to understand the effects of climate change, and to predict the high-risk areas of disease occurrence [6]. ENMs can be divided into “presence-absence” model and “presence-only” model according to the type of data used in model fitting. Many researchers have found that the presence-only approach is superior and preferable as only presence data is available in most cases [7].

As an infectious disease, the incidence of PTB is highly correlated with geographic information in the region and has obvious spatial and regional attributes. In this paper, we sought to analyze the epidemiologic features and identify the risk factors associated with spatial and temporal distribution for PTB using a maximum entropy ENM with Maxent version 3.4.1. The Maxent model was trained based on the locations of PTB occurrence and environmental factors that may affect the spatial distribution of PTB system. The predictive potential high-risk map was created by projecting the training model onto the research region with specific time. Furthermore, the performance of the model was assessed using the known locations of disease occurrence.

Materials And Methods

Study area

Nanjing, the capital city of Jiangsu province in eastern China between latitude 31°14"–32°37" and longitude 118°22"–119°14", was selected as the study area. It is one of the most important regional centre cities in the Yangtze River Delta economic zone,with a large population, strong mobility, complex natural and social environment, and unbalanced economic development. All of these factors may become the cause of a large-scale outbreak of PTB.

Data collection and management

In our study, disease occurrence data were collected from the Tuberculosis Management Information System, a subsystem of the China Information System for Disease Control and Prevention (CISDCP). Baidu and Google mapping services were used to search the residential address of patients and georeference the recorded cases. Overall, 2143 cases were selected from January 2017 to December 2017 to construct the training Maxent model, and 234 cases in May 2018 were collected to evaluate the prediction performance of the model.

Climatic variables with a resolution of 30 seconds (approximately 1 km), averaged across a temporal range of 1970 to 2000, were obtained from the WorldClim database (version 2.0) [8]. Elevation data were downloaded from the SRTM 90 m Digital Elevation Database, which were used to calculate the aspect, slope and composite topographic layers. The monthly maximum normalized difference vegetation index (NDVI) and land cover types were obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery. All the selected datasets were resampled to the same resolution, the same geographic extent, and constituted the environmental layers (Table 1).

Table 1
Environmental layers for model building and permutation importance

Variable	Description(unit)	Type	Data source	Permutation importance
prec	Precipitation(mm)	Continuous	http://www.worldclim.org, version 2.0	11.7%
tavg	Monthly average temperature(°C)	Continuous		20.3%
srad	Solar radiation(kJ m^{-2} day^{-1})	Continuous		31.1%
vapr	Monthly mean vapor pressure(kPa)	Continuous		31.8%
wind	Wind speed(m*s^{-1})	Continuous		0.0%
ndvi	Normalized difference vegetation index	Continuous	https://search.earthdata.nasa.gov/search	0.5%
lc	Land cover type	Categorical		1.4%
slope	Slope	Continuous	http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp	1.6%
aspect	Aspect	Continuous		1.6%

Epidemiological features analysis

Epidemiology [9, 10] is a science that studies the distribution and influencing factors of diseases, and then to explore the causes of disease, to formulate measures for disease preventing, controlling, and eliminating. Descriptive epidemiology [11] is one of the useful methods for epidemiological characteristics analysis, which is considered as the starting point of epidemiological research and the basis of other epidemiological research methods. This method was mainly adopted to describe the distribution, occurrence and development of diseases, and to describe the characteristics of some diseases in terms of time, place and population. In our research, we grouped the reported cases according to age, gender, register, occupation in each administrative division, and adopted the descriptive epidemiology method to analyze the epidemiological features.

Ecological niche model construction and evaluation

Maxent program was used to construct the training model for PTB in our study. It is a “presence-only” ENM based on maximum entropy algorithm [12], which is primarily used in spatial epidemiology to model species distributions using a set of environmental layers and a set of known occurrences [13]. Instead of predicting whether it exists or not, Maxent estimates the relative environmental adaptability of species existence [14].

“Auto features” was selected for Maxent running. The logistic output format with probability values ranging from 0 (unsuitable) to 1 (suitable) was selected for reflecting species presence. The model prediction was cross-validated by setting the random test and training ratio to 25:75. Default incidence was set to 0.5 as we were unable to get a reliable PTB incidence. In order to obtain a robust model, 10 models were created independently, and relevant outputs were averaged in our study. Maximum iterations were set to 5000, meanwhile, convergence threshold to $1*10^{-6}$. The evaluation data were combined with a random selection of 10000 background points. Other parameters were set by default.

A threshold-independent receiver operating characteristic approach (ROC) and area values under the curve (AUC) [15] were used to evaluate the ability of Maxent algorithm. This approach is typically used to evaluate the model by using a random 50% of occurrences calibration model and comparing the AUC with null expectations. The AUC quantifies the significance of the ROC. In general, AUC [16] values of 0.5 indicated that the model accuracy was no better than random, values of 1.0 indicated a perfect fitting, values above 0.70 were considered potentially useful. The spatial jackknife test was used to determine the percentage contribution of each environmental factors. This script can select the best model by evaluating the model’s omission rate, AUC, and model feature class complexity.

Model projection

A potential distribution map was created by projecting the training model onto the research regions with specific environmental factors. Meanwhile, the performance of the model was evaluated using the known PTB case data.

Results

Epidemiologic features of confirmed cases

Figure 1 listed the monthly number of PTB cases and average monthly cases in 2017. According to this study, the seasonal amount of PTB was not obvious. Except for the slight increase in the number of cases in March and September, the number of cases in other months was basically the same.

Table 2 listed the epidemiologic characteristics of PTB cases in Nanjing, China, 2017. The incidence of PTB cases was 25.9 per 100000 persons, of which 71.02% were males and 28.98% were females. According to the regional distribution of PTB cases, the annual incidence in Gulou, Jianye, Qinhuai regions was relatively low, while the annual incidence in Gaochun, Jiangning, Pukou, Liuhe, Lishui and Qixia regions was higher than the average incidence of Nanjing. Among them, the annual incidence of Gaochun is the highest, about 54.1 per 100000 persons, which is about 2.1 times the annual average incidence of Nanjing. Among PTB cases, the proportion of floating population was significantly higher than that of the permanent residents in developed regions. When the cases were classified by different age groups, the highest number of PTB patients was found in those aged 60 and above, followed by those aged 20 to 29,

and the lowest number was found in infants and children under 10 years of age. Classified with exact occupation, the highest number of PTB patients was found in farmers, followed by retirees, and the lowest number in teachers and medical staff.

Table 2
Epidemiologic characteristics of PTB cases in Nanjing, China, 2017.

Features	Xuanwu(%)	Qinhuai(%)	Jianye(%)	Gulou(%)	Pukou(%)	Qixia(%)	Yuhuatai(%)	Jiangning(%)	Liuhe(%)	Lishui(%
Cases[a]										
number	156(7.28)	199(9.29)	63(2.94)	147(6.86)	242(11.29)	182(8.49)	97(4.53)	405(18.90)	290(13.53)	129(6.0
incidence	$2.46*10^{-4}$	$1.98*10^{-4}$	$1.37*10^{-4}$	$1.18*10^{-4}$	$3.14*10^{-4}$	$2.63*10^{-4}$	$2.22*10^{-4}$	$3.33*10^{-4}$	$3.07*10^{-4}$	$2.98*10^{-4}$
Gender										
male	98(62.82)	140(70.35)	38(60.32)	112(76.19)	170(70.25)	126(69.23)	71(73.20)	285(70.37)	211(72.76)	98(75.9
female	58(37.18)	59(29.65)	25(39.68)	35(23.81)	72(29.75)	56(30.77)	26(26.80)	120(29.63)	79(27.24)	31(24.0
Register										
permanent	33(21.15)	51(25.63)	6(9.52)	73(49.66)	32(13.22)	38(20.88)	25(25.77)	247(60.99)	129(44.48)	68(52.7
floating	123(78.85)	148(74.37)	57(90.48)	74(50.34)	210(86.78)	144(79.12)	72(74.23)	158(39.01)	161(55.52)	61(47.2
Age										
0–9	0(0.00)	0(0.00)	0(0.00)	0(0.00)	0(0.00)	0(0.00)	0(0.00)	0(0.00)	1(0.34)	0(0.00)
10–19	6(3.85)	3(1.51)	1(1.59)	3(2.04)	3(1.24)	11(6.04)	5(5.15)	16(3.95)	11(3.70)	5(3.88)
20–29	40(25.64)	33(16.58)	12(19.05)	25(17.01)	62(25.62)	50(27.47)	26(26.80)	136(33.58)	54(18.62)	17(13.1
30–39	15(9.62)	24(12.06)	10(15.87)	16(10.88)	41(16.94)	19(10.44)	17(17.53)	59(14.57)	25(8.62)	11(8.53)
40–49	19(12.18)	18(9.05)	7(11.11)	20(13.61)	25(10.33)	19(10.44)	8(8.25)	38(9.38)	32(11.03)	15(11.6
50–59	20(12.82)	45(22.61)	11(17.46)	27(18.37)	39(16.12)	31(17.03)	14(14.43)	42(10.37)	54(18.62)	13(10.0
60–69	26(16.67)	36(18.09)	12(19.05)	28(19.05)	36(14.88)	25(13.74)	15(15.46)	55(13.58)	70(24.14)	31(24.0
70–79	17(10.90)	23(11.56)	8(12.70)	19(12.93)	26(10.74)	17(9.34)	7(7.22)	39(9.63)	30(10.34)	32(24.8
80–100	13(8.33)	17(8.54)	2(3.17)	9(6.12)	10(4.13)	10(5.49)	5(5.15)	20(4.94)	13(4.48)	5(3.88)
Occupation										
worker	21(13.46)	23(11.56)	3(4.76)	18(12.24)	33(13.64)	30(16.48)	15(15.46)	53(13.09)	25(8.62)	17(13.1
retiree	35(22.44)	53(26.63)	20(31.75)	49(33.33)	23(9.50)	27(14.84)	18(18.56)	38(9.38)	40(13.79)	4(3.10)
farmer	14(8.97)	28(14.07)	11(17.46)	16(10.88)	139(57.44)	46(25.27)	28(28.87)	132(32.59)	145(50.00)	95(73.6
teacher	1(0.64)	0(0.00)	1(1.59)	0(0.00)	4(1.65)	1(0.55)	1(1.03)	1(0.25)	0(0.00)	0(0.00)
student	18(11.54)	5(2.51)	1(1.59)	7(4.76)	7(2.89)	11(6.04)	6(6.19)	13(3.21)	9(3.10)	6(4.65)
doctor	3(1.92)	1(0.50)	1(1.59)	2(1.36)	2(0.83)	1(0.55)	1(1.03)	0(0.00)	0(0.00)	0(0.00)
server	6(3.85)	12(6.03)	2(3.17)	9(6.12)	18(7.44)	2(1.10)	5(5.15)	20(4.94)	5(1.72)	2(1.55)
others	58(37.18)	77(38.69)	24(38.10)	46(31.29)	16(6.61)	64(35.17)	23(23.71)	148(36.54)	66(22.75)	5(3.88)

% constituent ratio.

[a] Case incidence represents the incidence of cases calculated using cases number and total population in each administrative region, while others represents cases in each administrative region.

Variation of environmental factors

The relative importance of each environmental factor to the presence of PTB was shown in Fig. 2 and Table 1. The results indicated that the key environmental variables determining PTB occurrence include monthly mean vapor pressure, solar radiation, monthly average temperature, and precipitation. It can be seen from the response curves (Fig. 3) that how the predicted probability of presence changes, when each environmental factor changes. During this procedure, all other environmental variables were kept at their average sample value. When the logistic output probability of PTB presence was set to 0.5, the ideal conditions of the key environmental factors were (1) monthly average vapor pressure to be 1.71 kPa to 1.75 kPa, (2) solar radiation to be 18840 kJ m-2 day-1 to 19040 kJ m-2 day-1, (3) monthly average temperature to be 20.7 °C to 21.4 °C, and (4) monthly average precipitation to be 86 mm to 98 mm. All results indicated that the occurrence of PTB had a specific ecological niche with multi-dimensional environmental variables, playing an important role in the disease transmission cycle.

Statistical analysis and model evaluation

Figure 4 showed how testing, training omission, and predicted area vary with different cumulative thresholds. It can be seen from the figure that the omission of test samples closely matches predicted omission rate in the 10 models. Table 3 listed the AUCs of the training model and testing model, the omission rate and standard error of the AUC on the test data, as well as the p-value. All the AUCs for both training models and testing models were higher than 0.75. The average test AUC for the replicate runs was 0.822, with a standard deviation of 0.024. This indicated that both the testing model and the training model had good performance in predicting the potential high risk areas of PTB.

Table 3
Species Distribution Model (SDM) evaluation results.

Model Name	Training AUC	Testing AUC	Training omission rate	Testing omission rate	Standard Deviation	p-value
SDM1	0.834	0.841	0.012	0.016	0.025	< 0.0001
SDM2	0.832	0.853	0.004	0.016	0.023	< 0.0001
SDM3	0.833	0.850	0.004	0.000	0.023	< 0.0001
SDM4	0.838	0.799	0.009	0.031	0.031	< 0.0001
SDM5	0.836	0.833	0.011	0.047	0.028	< 0.0001
SDM6	0.839	0.781	0.005	0.063	0.030	< 0.0001
SDM7	0.837	0.793	0.007	0.048	0.029	< 0.0001
SDM8	0.835	0.834	0.005	0.000	0.023	< 0.0001
SDM9	0.836	0.810	0.007	0.016	0.028	< 0.0001
SDM10	0.835	0.822	0.009	0.000	0.024	< 0.0001

Potential risk areas of PTB

Figure 5 showed the predicted potential risk area, areas with reported PTB case, and heat map distribution of PTB in Nanjing in May 2018. The predicted potential risk area (Fig. 5A) was obtained by projecting the training model onto the research regions with various environmental factors in May 2018. The heat map of research regions (Fig. 5B) was calculated based on the known PTB cases in May 2018 using the kernel density estimation (KDE) method [17]. Without any prior knowledge and assumptions, the KDE method can be used to study the characteristics of case distribution and to measure the local intensity of a disease using the cases location itself. Our study indicated that the potential high-risk areas of Nanjing were mainly distributed in Gulou, Qinhuai, Xuanwu, and Jianye regions. Most of the predicted high-risk regions were confirmed by the known PTB cases.

Discussion

Spatial epidemiology, as a new branch of epidemiology, has been widely used in the research of PTB epidemiology [18, 19]. It uses geographic information technology [20] and spatial analysis technology [18] to describe and analyse the spatial distribution characteristics and development rules of diseases [21], to explore the determinants affecting population health [22], and to provide strategies and measures for disease prevention and health services [20]. Thanks to the development of geographic information technology, China has made remarkable achievements in using spatial data and spatial analysis methods to solve public health and health problems [21, 22, 23]. In this paper, we first analysed the epidemiological characteristics of PTB using the descriptive epidemiology method, and then used the Maxent ecological niche model to fit the spatial PTB occurrence data and environmental data to predict the key environmental factors and the potential high risk area for PTB in Nanjing.

Statistical analysis indicated that except for the small peak in March and September, PTB patients in Nanjing have no obvious seasonal phenomenon. The minor peak in March was mainly related to social factors, among which, employees returned to the city to work, bringing an increased flow of people, and thus leading to the spread of PTB [24]. The peak in September was mainly associated with natural factors. The possible reason for the small peak in September is the increasing number of physical examinations when the school year begins in September [25, 26, 27].

Regional economy directly affects the prevalence and control of PTB [28]. At present, Nanjing is in the period of rapid economic development, and the economic development of each district is unbalanced. In urban regions, the registration level of tuberculosis prevention and control institutions is higher than that in a suburban district, and people can get more employment opportunities, attracting a large influx of floating population (Fig. 6). Most of those floating population are farmers with relatively poor living standards and poor awareness of prevention. Another fact is that the male had much more PTB cases than female. This is mainly because a growing number of males have moved to cities to earn the family's bread, leaving women in rural areas to take care of their children. These people have a strong mobility and thus have many opportunities to be exposed to the source of infection.

One more phenomenon is that PTB cases can be found in all age groups. The number of PTB cases of infants and children under 10 years old are the lowest as they will be vaccinated with Bacillus Calmette-Guerin (BCG) promptly after birth [29]. Elderly people with low immunity levels are prone to direct infection or reignition of previously infected lesions, which is one of the most important reasons for the highest number of PTB cases in elders [30]. Young adults aged 20 to 29 have a relatively higher reporting rate. The number of floating workers in this age group is large, and their sense of self-protection is weak, which makes it easy for them to be infected with PTB.

Epidemiological surveillance indicated that environmental factors can affect the occurrence and development of PTB infectious disease. Temperature plays a fundamental role in the occurrence and development of PTB cases [31]. First, it directly influences the reproduction and transmission of vector-borne diseases [32]. Second, it affects the behaviors and ecological characteristics of hosts. Precipitation is also important meteorological conditions for the incidence of PTB cases. The higher the humidity, the longer the bacteria stays in the air, and the more likely it is to infect its host.

The potential risk map of PTB occurrence indicates that the economically developed central urban regions with much more floating populations are the high risk areas for the occurrence of PTB, which is consistent with actual case distribution and the heat map. However, there are many identified potential risk regions that have no reported PTB cases. For these regions with no reported PTB cases, it is mainly like that the PTB is a disease without specific clinical manifestations and it is easily misdiagnosed and overlooked by people. Another possible reason is that most of these areas are less developed regions with less experienced medical staff, thus reliable diagnostics are difficult to obtain. Therefore, epidemiological surveillance should be enhanced in these areas.

Conclusions

The prediction of the potential distribution of PTB disease is a much more complex work. Our research has successfully provided an accurate ENM model for predicting the potential geographical distribution of PTB disease, taking into account the habitat adaptability of species. In our research, we only took environmental factors into consideration although the incidence of PTB is affected by economy, geographical environment, climate, health services and other factors. Thus, to validate the prediction result of the model, surveillance of PTB cases should be enhanced.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee at the second hospital of Nanjing, Nanjing University of Chinese Medicine. The data about PTB cases were downloaded from existing online databases after elected consent by the local ethical committee. Meanwhile, no sample of human and animal was included, and all data were planned to be analyzed anonymously. All aspects of the study are complied with the Declaration of Helsinki.

Consent to publish

Not applicable.

Availability of data and materials

All relevant data are within the manuscript and its Supplementary Information files. S1 Data: Reported PTB cases used in our research; S2 Data: Population and GDP information in each administrative division of our research area; S3 Data: Environmental data set in our research.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the scientific research project of Jiangsu Commission of Health (NO.Y2018001) and the National Nature Science Foundation of China (NO. 81601794). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

HYZ and HYY designed the study. HYZ wrote the manuscript and analysed model results. CKS contributed to spatial analysis and Maxent modelling. WM analysed epidemiologic data. JYS conducted fields sampling and collated the data throughout this study. XZ analysed environmental data. GPY contributed to interpreting the results. All authors read and approved the final manuscript.

Acknowledgements

We thank Steven Phillips, from AT&T Labs-Research, Princeton University, for providing Maxent model. We are very grateful Jianli Hu from Jiangsu Provincial Centre for Disease Control and Prevention for revising the final manuscript.

Abbreviations

PTB: Pulmonary tuberculosis; ENM: ecological niche model; SRTM: Shuttle Rader Topography Mission; NDVI: normalized difference vegetation index; MODIS: Moderate Resolution Imaging Spectroradiometer; ROC: receiver operating characteristic curve; AUC: area under the curve; KDE: kernel density estimation.

References

1.
Zhao MX, Wang T, Yuan XY, Du WM, Lin MX, Shen YB. Comparison of minocycline and azithromycin for the treatment of mild scrub typhus in northern China. Int J Antimicrob Agents. 2016;48:317–20.
2.
Li XX, Wang LX, Zhang H, Jiang SW, Fang Q, Chen JX, Zhou XN. Spatial variations of pulmonary tuberculosis prevalence co-impacted by socio-economic and geographic factors in People's Republic of China, 2010. BMC Public Health. 2014;14:257. DOI:10.1186/1471-2458-14-257.
3.
Maciel ELN, Pan W, Dietze R, Peres RL, Vinhas SA, Ribeiro FK, Palaci M, Rodrigues RR, Zandonade E, Golub JE. Spatial patterns of pulmonary tuberculosis incidence and their relationship to socio-economic status in Vitoria, Brazil. Int J Tuberc Lung Dis. 2010;14(11):1395–402(8).
4.
Xu M, Cao CX, Wang DC, Kan B. Identifying Environmental Risk Factors of Cholera in a Coastal Area with Geospatial Technologies. Int J Environ Res Public Health. 2014;12(1):354–70.
5.
Brotons L, Thuiller W, Araujo MB, Alexandre HH. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography. 2004;27(4):437–48.
6.
Yu HY, Sun CK, Liu WD, Li ZF, Tan ZM, Wang XC, Hu JL, Shi SQ, Bao CJ. Scrub typhus in Jiangsu Province, China: epidemiologic features and spatial risk analysis. BMC Infect Dis. 2018;18:372. doi:10.1186/s12879-01803271-x.
7.
Sean PM, Christophe FR, Antoine G, John MD. Pattern-recognition ecological niche models fit to presence-only and presence–absence data. Methods Ecol Evol. 2014;5(8):761–70.
8.
Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37(12):4302–15.
9.
Li YL, Li XL, Liang S, Fang LQ, Cao WC. Epidemiological features and risk factors associated with the spatial and temporal distribution of human brucellosis in China. BMC Infect Dis. 2013;13:547.
10.
Wu L, Qiao FY, Li YH, Zhu YF, Liu WD. Epidemiological characteristics and spatial-temporal clustering analysis on tuberculosis in China, from 2004 to 2012. Jiangsu J Prev Med. 2014;25:19–22.
11.
Golden J, Barry WE, Jang G, Nguyen N, Bliss D. Pediatric Morgagni diaphragmatic hernia: a descriptive study. Pediatr Surg Int. 2017;33:771–5. doi:10.1007/s00383-017-4078-3.
12.
Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol model. 2006;190:231–59.
13.
Phillips SJ, Dudik M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 2008;31:161–75.
14.
Bui CM, Gardner L, MacIntyre R, Sarkar S. Influenza. A H5N1 and H7N9 in China: A spatial risk analysis. PLoS ONE. 2017;12(4):e0176903.
15.
Fielding AH, Bell JFA. review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv. 1997;24(1):38–49.
16.
Peterson AT, Papes M, Soberon J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecol model. 2008;213:63–72.
17.
Chen YC. A tutorial on kernel density estimation and recent advances. Biostatistics Epidemiology. 2007;1(1):161–87.
18.
Harling G, Castro MC. A spatial analysis of social and economic determinants of tuberculosis in Brazil. Health Place. 2014;25:56–67.
19.
Lai PC, Low CT, Tse WS, Tsui CK, Lee H, Hui PK. Risk of tuberculosis in high-rise and high density dwellings: An exploratory spatial analysis. Environ Pollut. 2013;183:40–5.
20.
Russell SK, Eric D, Jan ME. Advances in spatial epidemiology and geographic information systems. Ann Epidemiol. 2017;27(1):1–9.
21.
Yang K, Sun LP, Liang YS, Wu F, Li W, Zhang JF, Huang YX, Hang DR, Liang S, Bergquist R, Zhou XN. Schistosoma japonicum risk in Jiangsu province, People's Republic of China: identification of a spatio-temporal risk pattern along the Yangtze River. Geospat Health. 2013;8(1):133–42.
22.

Liu WD, Zhu YF, Qi X, Xu K, Ge AH, Ji H, Ai J, Bao CJ, Tang FY, Zhou MH. Risk assessment on the epidemics of human infection with a novel avian influenza A (H7N9) virus in Jiangsu Province, China. J Biomed Mater Res. 2013;27(3):163–6.
23.
Zhang WY, Fang LQ, Jiang JF, Hui FM, Glass GE, Yan L, Xu YF, Zhao WJ, Yang H, Liu W, Cao WC. Predicting the Risk of Hantavirus Infection in Beijing, People's Republic of China. Am J Trop Med Hyg. 2009;80(4):678–83.
24.
Tang L, Su Q. Epidemiological characteristics of tuberculosis in Huaian City, 2005–2014. Modern Preventive Medicine. 2016;43(5):780–1,809.
25.
Zhang X, Huang WX. Epidemiological analysis of tuberculosis in Tujia nationality area from 1998 to 2003. Chinese Journal of Practice Medicine. 2004;3(12):1126–7.
26.
Wang R, Li C, Zhao P. Epidemiological characteristics of pulmonary tuberculosis among transient and local population in Nanjing, 2010. Jiangsu J Prev Med. 2013;24(2):12–4.
27.
Zhou BB, Lu HC, Gong DG, Xu GZ. Epidemiological characteristic of student cases with pulmonary tuberculosis in Yinzhou district of Ningbo city 2005–2009. Chinese Journal of School Health. 2011;32(5):599–600.
28.
Schepisi MS, Gualano G, Piselli P, Mazza M, Angelo DD, Fasciani F, Barbieri A, Rocca G, Gnolfo F, Olivani P, Ferrarese M, Codecasa LR, Palmieri F, Girardi E. Active tuberculosis case finding interventions among immigrants, refugees and asylum seekers in Italy. Infect Dis Rep. 2016;8(2):6594.
29.
Soysal A, Millington KA, Bakir M, Dosanjh D, Aslan Y, Deeks JJ, Efe S, Staveley I, Ewer K, Lalvani A. Effect of BCG vaccination on risk of Mycobacterium tuberculosis infection in children with household tuberculosis contact: a prospective community-based study. Lancet. 2005;366(9495):1443–51.
30.
Lee CH, Wang JY, Lin HC, Lin PY, Chang JH, Suk CW, Lee LN, Lan CC, Bai KJ. Treatment delay and fatal outcomes of pulmonary tuberculosis in advanced age: a retrospective nationwide cohort study. BMC Infect Dis. 2017;17:449.
31.
Li XX, Wang LX, Zhang J, Liu YX, Zhang H, Jiang SW, Chen JX, Zhou XN. Exploration of ecological factors related to the spatial heterogeneity of tuberculosis prevalence in P. R. China. Glob Health Action 2014, doi: 10:3402/gha.v7.23620.
32.
Yano Y, Shiraishi S, Uchida TA. Effects of temperature on development and growth in the tick, Haemaphysalis longicornis. Exp Appl Acarol. 1987;3(1):73–8.

Figures

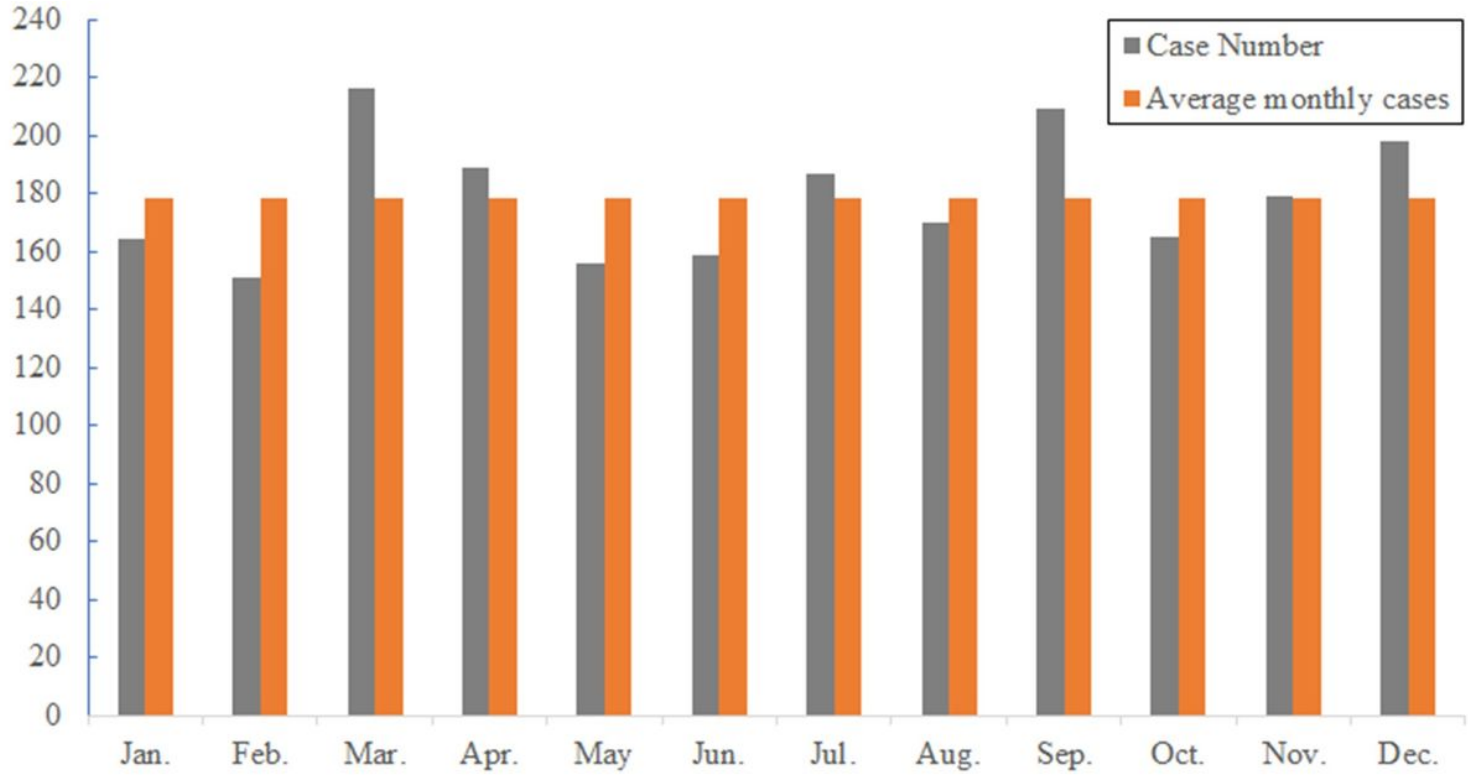

Figure 1

Registration of PTB in Nanjing, China, 2017.

Figure 2

Jackknife analysis results of training gain, test gain, and AUC. Average values were shown. (A) Jackknife of regularized training gain for PTB. (B) Jackknife of test gain for PTB. (C) Jackknife of AUC for PTB).

Figure 3

Response curves for the variables related to presence of PTB. Red lines were mean response of the 10 replicate Maxent runs and blue bars represent the mean +/- one standard deviation. (A) Response curve for precipitation. (B) Response curve for solar radiation. (C) Response curve for monthly mean vapor. (D) Response curve for monthly average temperature. (E) Response curve for land cover type. (F) Response curve for normalized difference vegetation index. (G) Response curve for wind speed. (H) Response curve for slope. (I) Response curve for aspect.

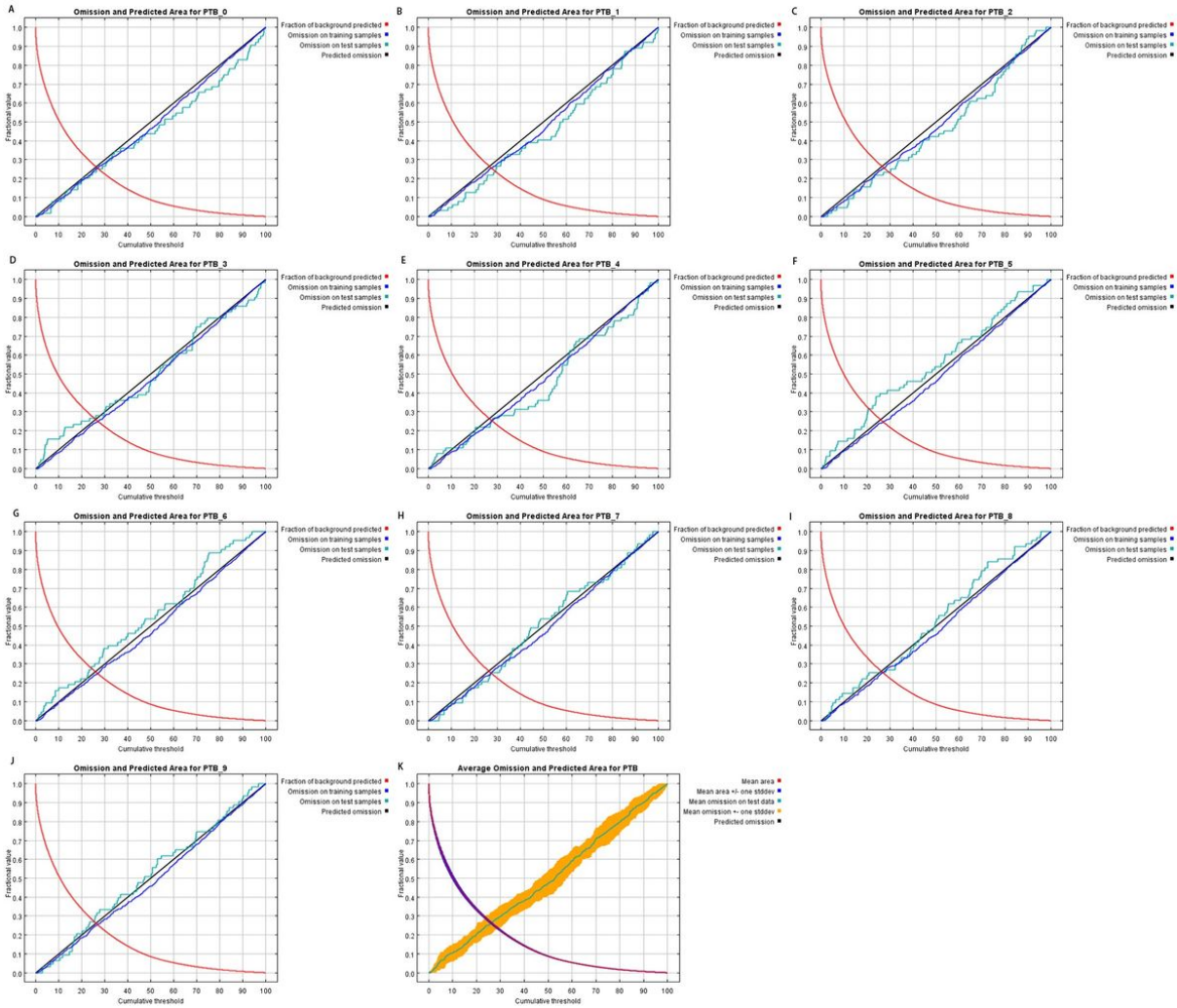

Figure 4

Omission and predicted areas for (A) PTB_0, (B) PTB_1, (C) PTB_2, (D) PTB_3, (E) PTB_4, (F) PTB_5, (G) PTB_6, (H) PTB_7, (I) PTB_8, (J) PTB_9, and (K) average omission and predicted area for PTB.

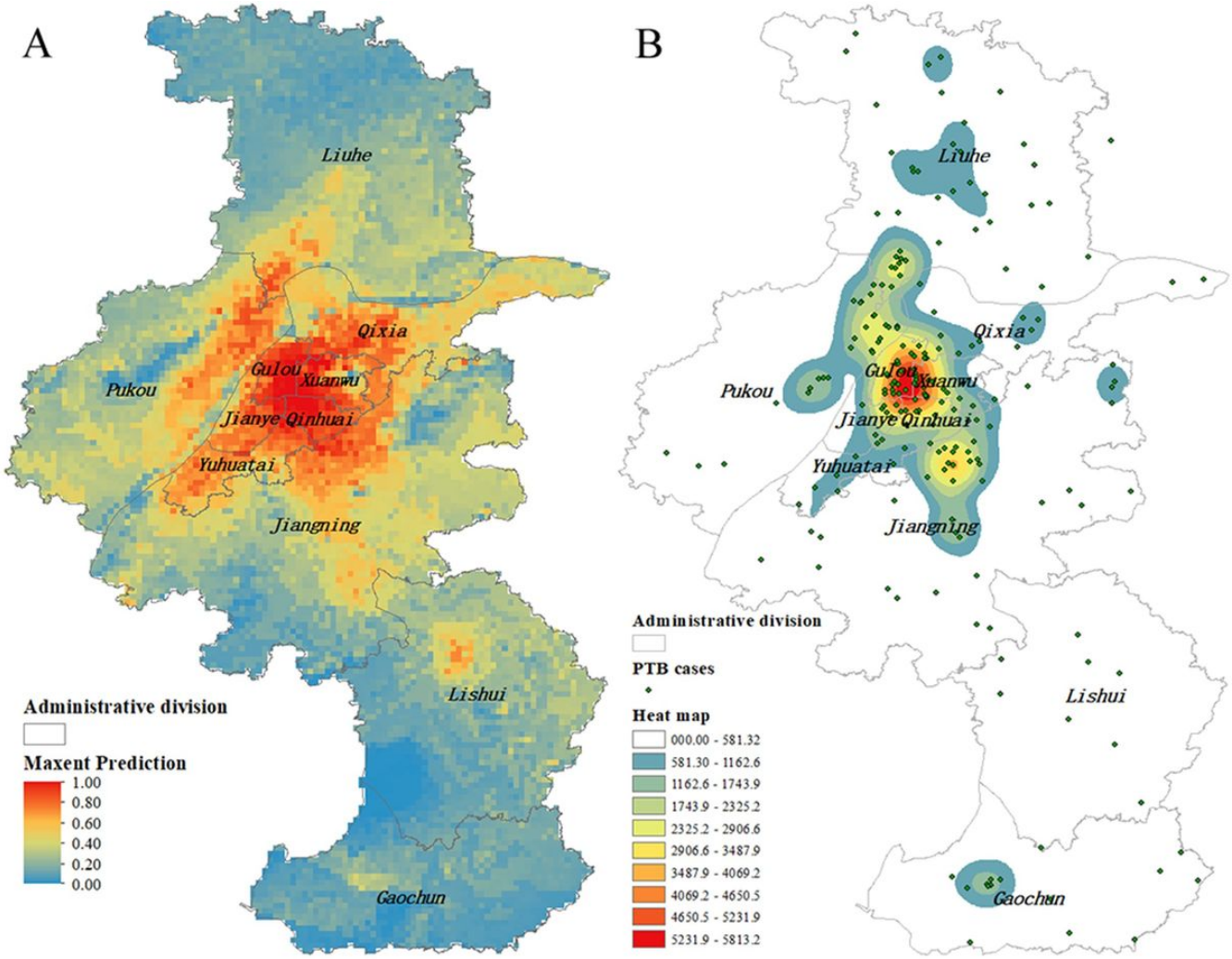

Figure 5

Distribution of PTB diseases in May, 2018, Nanjing, China. (A) Predicted potential high risk areas, (B) reported PTB occurrence and heat map. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Figure 6

Population information of Nanjing in 2018. (This information can be collected from statistical yearbook of Nanjing)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- wind.asc
- prec.asc
- ndvi.asc
- slope.asc
- lc.asc
- aspect.asc
- srad.asc
- tavg.asc
- S2.xlsx
- vapr.asc

- S1.xlsx