Spatial modeling of malaria incidence rates in Sistan and Baluchistan province, Islamic Republic of Iran

Masoud Salehi, PhD, Student, *Kazem Mohammad,* PhD, *Mahmud M. Farahani,* PhD, *Hojjat Zeraati,* PhD, *Keramat Nourijelyani,* PhD, *Farid Zayeri,* PhD.

ABSTRACT

الأهداف: تحديد أثر العوامل البيئية على خطورة الملاريا وتصور الخريطة الحيزية لمعدلات الحدوث القياسية للملاريا في مقاطعة سيستان وبلوشستان في الجمهورية الإيرانية.

الطريقة: أجريت دراسة مقطعية، تم فيها دراسة البيانات لـ42162 حالة ملاريا مسجلة جديدة خلال الفترة مابين 21 مارس 2001م وحتى 21 مارس 2006م. من أجل وصف الإتحاد الإحصائي بين العوامل البيئية وخطورة الملاريا. تم استخدام طريقة الطراز العام للخط المختلط، بالإضافة إلى استخدام الطلب الثاني المستقر وتخطيط الدوالي من أجل تحديد التركيبة المتصلة بالحيز المناسب بين المعدلات القياسية لحدوث الملاريا، ولتقديم خريطة خطر الملاريا المناسبة في منطقة الدراسة.

النتائج: كشفت النتائج التي حصل عليها من النموذج الحيزي إن الرطوبة (p=0.0004)، درجة الحرارة (p<0.0001)، والارتفاع (p<0.0001) كانت موجبة ومترسبة (p=0.0029) ومتصلة عكسياً مع خطورة الملاريا. إضافتاً إلى ذلك، أظهرت خريطة خطر الملاريا المبنية على القيم المتنبئة أن الجزء الجنوبي من هذه المقاطعة (بلوشستان) لديها خطورة أعلى من الملاريا مقارنة مع المناطق الشمالية (سيستان).

خاتمة: نظراً لكون العوامل البيئية المؤثرة في خطورة الملاريا خارجة عن سيطرة الإنسان، لذا يجب إعارة المزيد من الانتباه لهذه المقاطعة من قبل القائمين على الأنظمة الصحية، إضافة إلى المناطق الأخرى التي ترتفع بها درجة الحرارة، ونسبة الرطوبة وكذلك انخفاض هطول الأمطار.

Objectives: To identify the effect of environmental factors on malaria risk, and to visualize spatial map of malaria standard incidence rates in Sistan and Baluchistan province, Islamic Republic of Iran.

Methods: In this cross-sectional study, the data from 42,162 registered new malaria cases from 21 March 2001 (Iranian new year) to 21 of March 2006 were studied. To describe the statistical association between environmental factors and malaria risk, a generalized linear mixed model approach was utilized. In addition, we used the second ordered stationary Kriging, and a variogram to determine the appropriate spatial correlation structure among the malaria standard incidence rates, and provide a proper malaria risk map in the area under study.

Results: The obtained results from the spatial modeling revealed that humidity (p=0.0004), temperature (p<0.0001), and elevation (p<0.0001) were positively, and precipitation (p=0.0029) was inversely correlated with the malaria risk. Moreover, the malaria risk map based on the predicted values showed that the south part of this province (Baluchistan), has a higher risk of malaria, compared to the northern area (Sistan).

Conclusion: Since the effective environmental factors on malaria risk are out of human's control, the health policy makers in this province should pay more attention to the areas with high temperature, elevation, and humidity, as well as, low rainfall districts.

Saudi Med J 2008; Vol. 29 (12): 1791-1796

From the Department of Biostatistics and Epidemiology (Salehi, Mohammad, Farahani, Zeraati, Nourijelyani), School of Public Health, Tehran University of Medical Sciences, and the Department of Biostatistics (Zayeri), Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Received 13th July 2008. Accepted 8th November 2008.

Address correspondence and reprint request to: Professor Kazem Mohammad, Department of Biostatistics and Epidemiology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. Tel. +98 (21) 88973901. Fax. +98 (21) 88974462. E-mail: mohamadk@tums.ac.ir / Salehi74@yahoo.com

Malaria is one of the most important heath problems that affect over 100 million persons, and kills approximately one million people, especially in children under 5 years.[1] This infectious communicable disease is more prevalent in semiarid areas such as the Middle East, and African countries. Therefore, measuring malaria burden, and determining its related factors in a population, is a challenge in most developing countries. The malaria incidence rate is strongly associated with factors such as massive displacements, famine, lack of efficient disease surveillance, and environmental conditions.[2,3] Such as other Middle East countries, malaria is an important health problem in Iran, particularly in the eastern and south-eastern areas. According to the report of the Malaria Office of the Ministry of Health of Iran, the annual malaria cases were at least 15000 during 21 of March 2001 (Iranian new year), and 21 of March 2006.[4] Among these cases, approximately 60% were reported from a vast province in the east (Sistan), and southeast of Iran (Baluchistan).[5] The analysis of the geographical distribution of the incidence of the disease, and its relationship to potential risk factors, has an important role to play in various kinds of public health and epidemiological studies. In this context, using proper statistical models is a helpful strategy in the analysis of geographical data, and describing the statistical relationship between potential risk factors and disease incidence/prevalence.[6,7] Additionally, spatial modeling is a useful statistical tool for producing a proper incidence/prevalence map of the area illustrating the variation in malaria risks.[8-15] Many standard statistical models assume independence of observations. However, the majority of infectious diseases, such as malaria cases, cluster due to underlying common environments. When spatially correlated data are analyzed, this independent assumption leads to overestimation of the statistical significance of covariates.[16] Spatial models incorporate the spatial correlation according to the way the geographical information is available. For rate or count data, the spatial correlation is defined by a neighborhood structure. For geo-statistical data, the spatial correlation is usually considered as a function of the distance between locations. In the malaria rates, due to its strong spatial correlation, the regression modeling results should be adjusted using convenient generalized linear mixed models (GLMMs), and variograms. Recently, Bithell,[17] Diggle,[18] Lawson,[19,] and Lawson et al [20] provided brilliant reviews of disease mapping using spatial modeling. Despite the high incidence rate of malaria in Sistan and the Baluchistan province (SBP), we found no published article on spatial modeling of the incidence rate of malaria and its related factors in this part of Iran. In this study, therefore, we aim to explore the effective environmental factors of malaria such as precipitation, humidity, elevation, and temperature in SBP, Islamic Republic of Iran.

Methods. This epidemiologic cross-sectional research was carried out in SBP in southeastern Iran as a part of a PhD dissertation in the Department of Epidemiology and Biostatistics of Tehran University of Medical Sciences, Tehran, Iran. This study was approved by the Ethics Committee of Tehran University of Medical Sciences.

Study area. Covering a surface area of 181,470.9 square kilometers, Sistan and Baluchistan province is twice the size of England. Sistan and Baluchistan provinces consist of 2 main parts. Sistan, in the northern part, is in the neighborhood of Afghanistan, and Baluchistan, in the southern part, is in the neighborhood of Pakistan and Oman Sea. Sistan and Baluchistan has 10 provinces (Zahedan, Zabol, Chabahar, Iranshahr, Saravan, Khash, Konarak, Zahak, Sarbaz and Nikshahr), 37 districts, and 97 subdistricts. Zahedan is the capital province. Khash, the highest city is situated 1,394 meters above sea level, and Chabahar, is 7 meters above sea level, and is the lowest city of the province. Sistan and Baluchistan provinces' has a common border of 1,265 km with southeastern neighboring countries.

Data source. In this study, we used registered malaria data from urban and rural health centers in SBP. Variables such as patients' and their parents' names, age, gender, type of accommodation (permanent or temporary), nationality (Iranian or immigrant), location (name of village or city), and the date of obtaining and reading blood films were gathered from 21 of March 2001 (Iranian new year) to 21 of March 2006. A MicroSoft-Access data bank was utilized to locate the recurrent malaria cases. In the Islamic Republic of Iran, the national census is conducted every 10 years by the Statistical Center of Iran (2 last censuses were in 1996 and 2006). The adjusted midyear population size for the period between the 2 censuses is calculated, and represented by the Statistical Center of Iran.[21] We used these adjusted midyear population size for calculating the incidence rates of malaria. Overall, the information from 37 districts was collected. For each district a standard incidence rate (SIR) was calculated for the described period. To calculate these SIRs, it

*The full text including Appendix is available in PDF format on Saudi Medical Journal website (www.smj.org.sa)

was assumed that the entire population of a district was exposed to the risk of malaria during this period, that is, each person contributed exactly one person-year of exposure. In computing the malaria SIRs, the repeated episodes of the subject were considered just as an observed subject for each year. The environmental explanatory variables such as monthly precipitation, average of monthly humidity, elevation, and average of maximum daily temperature in a month for the above mentioned period, were obtained from the registered climate data in Islamic Republic of Iran Meteorological Organization (IRIMO).[22] By using Geographical Information System (GIS), the average value of each variable for each district under study was calculated by averaging the pixel variable over the area in the district.[23]

Statistical analysis. For the epidemiological description of the data, the SIR of malaria was computed separately, for the total population and by gender, in different districts of SBP.[24] The spatial correlation between district and malaria SIRs, was assessed using Moran's I statistics.[16] This analysis showed that the malaria SIR are severely spatially correlated. In addition, since we had unobserved malaria cases in some areas, the Kriging approach was used for predicting these unobserved values. In the next step, a variogram was utilized to find the appropriate autocorrelation structure (sill, nugget and range).[16,25] Finally, a spatial generalized linear mixed modeling approach was utilized for describing the relationship between environmental and climatic explanatory variables, and malaria SIRs. Appendix A shows more details on the structure of the data, and the utilized model for data analysis. We used R?? codes for computing the variogram and the SAS software, version 9.1, for spatial modeling of the data. The interested reader can refer to Appendix B for R?? codes. ArcGIS software, version 9, was also used to visualize the geographical map of malaria in SBP.

Results. In general, a 42,162 malaria new cases were registered from March 2001 to March 2006 in SBP. Among them, 27700 cases (65.7%) were male, and 14461 cases (34.3%) were female. This study also showed that 2530 (6%) of these patients were children less than 5 years old, and 35078 cases (83.2%) were adults older than 15 years. Plasmodium vivax was observed in 33097 (78.5%), Plasmodium falciparum in 8601 (20.4%), and mixed infections were observed in 464 (1.1%). Among the detected new cases, 3837 cases (9.1%) were from the urban, and 38,325 cases (90.9%) were from the rural areas. **Table 1** shows more description on the population and malaria SIRs (**Table 1**). Malaria SIRs showed significant positive spatial correlation with Moran's I statistics (p<0.001). To describe the relationship between climatic and environmental covariates and observed malaria SIRs while accounting for spatial correlation of the data, an iterative approach of variogram and GLMMs was utilized. **Figure 1** displays the variogram of deviance residuals of the final model of malaria SIRs in SBP. Regarding the presented variogram in **Figure 1**, the GLMM model with convenient correlation structure for this data set was fitted. **Table 2** shows the obtained results. According to the presented results in **Table 2**, one can conclude that temperature

Table 1 - Malaria SIR in Sistan and the Baluchistan province between March of 2001 and March of 2005.

Year	Gender	Total population	Malaria new cases	Age mean±SD	SIR	95% CI
2001	Male	1011570	4103	24.2±18.07	0.041	0.039 – 0.044
	Female	1005606	1759	23.7±17.64	0.018	0.017 – 0.019
	Total	2017176	5862	23.9±17.98	0.029	0.028 – 0.031
2002	Male	1045081	3733	22.7±16.93	0.036	0.034 – 0.038
	Female	1038921	1758	21.9±18.76	0.017	0.016 – 0.018
	Total	2084002	5491	22.2±17.35	0.027	0.025 – 0.028
2003	Male	1078172	9935	22.9±17.16	0.093	0.088 – 0.099
	Female	1071827	4464	21.9±17.49	0.042	0.040 – 0.045
	Total	2150009	14399	22.5±17.28	0.067	0.064 – 0.072
2004	Male	1112977	4178	23.4±16.29	0.038	0.036 – 0.040
	Female	1106416	2058	22.7±17.79	0.019	0.018 – 0.020
	Total	2219393	6236	23.2±16.83	0.028	0.027 – 0.030
2005	Male	1148423	6308	23.5±16.21	0.055	0.053 – 0.059
	Female	1141653	3866	24.0±18.31	0.034	0.032 – 0.037
	Total	2290076	10174	23.6±16.96	0.045	0.042 – 0.047

SD - standard deviation, SIR - standard incidence rate, CI - confidence interval

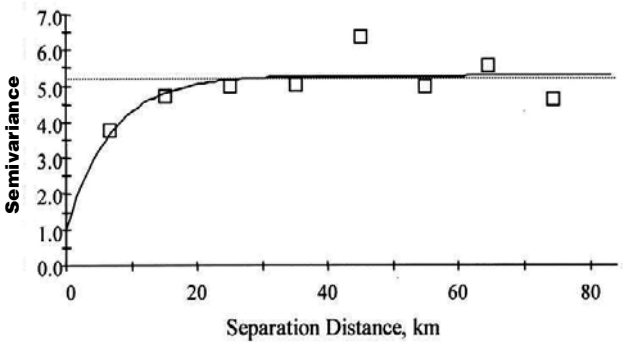

Figure 1 - Variogram of deviance residuals of the model of malaria standards incidence rates for the population of the districts of Sistan and Baluchistan province, Islamic Republic of Iran.

Table 2 - Spatial modeling results for assessing the environmental and climatic factors on malaria standard incidence rates.

Explanatory factor	Est	SE	*P*-value	95% CI
Precipitation	-0.0105	0.0033	0.0029	-0.0171 – -0.0039
Temperature	0.4629	0.0621	<0.0001	0.3364 – 0.5894
Elevation	0.0033	0.0005	<0.0001	0.0023 – 0.0043
Humidity	0.0403	0.0101	0.0004	0.0197 – 0.0609

Est - estimate of the model parameter, SE - standard error of the estimate, CI - confidence interval for the estimate

(p<0.0001), humidity (p=0.0004), elevation (p<0.0001) and precipitation (p=0.0029) had a significant effect on malaria SIR in SBP. In addition, the sign of the estimates tells us that temperature, humidity and elevation had a direct effect, and precipitation has an inverse effect on malaria incidence rates. In the next step, we provided the geographical map of malaria in SBP based on the predicted SIR values by the described GLMM. Our findings show that Zabol, Zahak, and Khash had the lowest, and Nikshahr, Sarbaz, and Iranshahr had the highest malaria SIRs among the SBP counties. In general, these results show that Baluchistan has a considerably higher malaria incidence rate compared to Sistan.

Discussion. Generally, transmission of malaria occurs at 640 North and 320 South of the earth, where there are favorable conditions for the life cycle of malaria parasite.[26] Sistan and Baluchistan provinces are also located in this limited area. Accurate maps of malaria incidence are important tools in malaria control as they can guide interventions, and assess their effectiveness. These maps rely on predictions of risk at locations without observed prevalence data. Malaria is an environmental disease, and environmental factors mainly temperature, rainfall, and humidity are good predictors of transmission, but the relation between these environmental factors, mosquito abundance, and malaria incidence/prevalence is not linear. This relation can be established only by means of adequate spatial statistical models. By determining the relations between malaria and the environment, the burden of malaria can be estimated at places where data on transmission are not available, and high-risk areas can be identified. Accurate maps of malaria transmission can guide intervention strategies, and thus optimize the use of limited and financial resources to areas of utmost need. In addition, early warning systems can be developed to predict epidemics of malaria from environmental changes.[27]

While there is a growing body of literature on the use of spatial modeling for malaria research and control, there have been a few issues on malaria indexes in Iran,[28,29] and there is no comprehensive survey and published article on malaria incidence, or prevalence rate in SBP. In this study, we used a GLMM to assess the relationship between malaria SIRs and environmental factors and provide an accurate malaria incidence map in SBP. Elevation has long been recognized to be associated with malaria due to its association with cooler temperatures and humidity.[30-32] In the present study, our findings revealed that the elevation had a positive association with malaria SIRs. Due to the large variation in elevation distribution in SBP (from 7 meters to 1394 meters above the sea level), this factor provides a useful measure of transmission risk, and malaria incidence rate. A number of research has found a strong correlation between the malaria incidence rate and variations in the environmental variables during several preceding months,[33] or with inter-annual variations in these variables.[34] In many studies, humidity, temperature, and rainfall are considered major risk factors that affect the life cycle and breeding of mosquitoes.[35] The model derived in the present study uses monthly average of maximum daily temperature, humidity, and total monthly precipitation. This model showed a strong correlation between malaria SIRs and precipitation. We know that reasonable amounts of rainfall create additional breeding sites for mosquitoes, thereby, increasing its population. Thus, we expect a positive association between malaria SIRs in the region and precipitation, however, in our survey we found a negative correlation between precipitation and malaria risk. This negative correlation can be explained with huge amounts of rain that wash the ground and kill the eggs.[35]

In addition, our statistical modeling revealed a positive

effect of temperature and humidity on malaria SIRs. These findings confirm the obtained results from other studies in different countries.[27,34,36] Factors related with population vulnerability are also critically important in malaria transmission. The presence of parasite resistance to the usual anti-malaria and to insecticides, population movements, and the presence of other underlying infections (namely, HIV) are responsible for a large part of the variability in the incidence of malaria. Considering all these factors, in addition to environmental variables in statistical models, to predict malaria incidence is complex, and not yet well understood. This has led some researchers to develop models, in which incidence rates are standardized with respect to non-climatic variables, so that the influence of climate on fluctuations in the malaria rate can be seen more clearly.[33,37] In this study, we also used the described model to obtain the predicted values of SIRs in different districts of SBP. Differences between the observed and predicted malaria SIRs, are highly due to other above-mentioned effective factors, which were not considered in our statistical modeling. Lack of information on some important factors such as patients' educational level, socio-economic status, occupation, traditional bedroom type was the major limitation of this study.

In general, the present study showed significant association between environmental factors and malaria risk in SBP. Since these environmental factors are out of human's control, the health policy makers in this province should pay more attention to the areas with high temperature, elevation, and humidity, as well as, low rainfall districts.

References

1. WHO/UNISEF. The World Malaria Report 2005. Geneva (Switzerland): World Health Organization; 2005.
2. Nájera JA, Kouznetzsov RL, Delacollette C, editors. Malaria Epidemics: Detection and Control, Forecasting and Prevention. WHO/MAL/98.1084. Geneva (Switzerland): World Health Organization; 1998.
3. Kazembe LN, Kleinschmidt I, Sharp BL. Patterns of malaria–related hospital admissions and mortality among Malawian children: an example of spatial modeling of hospital registered data. ***Malar J*** 2006; 5: 93.
4. Ministry of Health and Medical Education of Iran. Malaria Center. [update 2008 Oct 7]. Available at URL:http://www.mohme.gov.ir/health/index.htm
5. Islamic Republic of Iran. Overview of malaria control activities and program progress. World Health Organization. [cited 2005 April 27]. Available at URL: http://www.rbm.who.int/wmr2005/profile/iran.pdf
6. Gething PW, Noor AM, Gikandi PW, Ogara EA, Hay SI, Nixon MS, et al. Improving imperfect data from health management information systems in Africa using space time geo-statistics. ***PLoS Med*** 2006; 3: 271.
7. Elliot P, Wakefield J, Best N, Briggs DJ, editors. Spatial epidemiology-methods and applications. Oxford (UK): Oxford University Press; 2000.
8. Kleinschmidt I. Spatial statistical analysis, modeling and mapping of malaria in Africa [Thesis]. Basel (Switzerland): University of Basel: Faculty of Philosophy and Natural Sciences; 2001.
9. Gemperli A. Development of spatial statistical methods for modeling point-referenced spatial data in malaria epidemiology [Thesis]. Basel (Switzerland): University of Basel: Faculty of Philosophy and Natural Sciences; 2003.
10. Craig MH, Snow RW, Le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. ***Parasitol Today*** 1999; 15: 105-111.
11. Thomson MC, Connor SJ, D'Alessandro U, Rowlingson B, Diggle P, Cresswell M, et al. Predicting malaria infection in Gambia children from satellite data and bed net use surveys: the importance of spatial correlation in the interpretation of results. ***Am J Trop Med Hyg*** 1999; 61: 2-8.
12. Hay SI, Omumbo JA, Craig MH, Snow RW. Earth observation, geographic information system and Plasmodium falcipyrum malaria in sub-Saharan Africa. ***Adv Parasitol*** 2000; 47: 173-215.
13. Kleinschmidt I, Bagayoko M, Clarke GPY, Craig M, Le Sueur D. A spatial statistical approach to malaria mapping. ***Int J Epidemiol*** 2000; 29: 355-361.
14. Rogers DJ, Randolph SE, Snow RW, Hay SI. Satellite imagery in the study and forecast of malaria. ***Nature*** 2002; 415: 710-715.
15. Omumbo JA, Hay SI, Goetz SJ, Snow RW, Rogers DH. Updating historical maps of malaria transmission intensity in East Africa using remote sensing. ***Photogramm Eng Rem Sens*** 2002; 68: 161-166.
16. Cressie NAC, editor. Statistics for spatial data. 1st ed. New York (NY): John Wiley & Sons Inc; 1993.
17. Bithell JF. A classification of disease mapping methods. ***Stat Med*** 2000; 19: 2203-2215.
18. Diggle PJ. Overview of statistical methods for disease mapping and its relationship to cluster detection. In: Elliot P, Wakefield JC, Best NG, Briggs DJ, editors. Spatial epidemiology: methods and applications. Oxford (UK): Oxford University Press; 2000.
19. Lawson AB. Tutorial in biostatistics: disease map reconstruction. ***Stat Med*** 2001; 20: 2183-2203.
20. Lawson AB, Biggeri A, Böhning D, Lesaffre E, Viel JF, Clark A, et al. Disease mapping models: an empirical evaluation. ***Stat Med*** 2000; 19: 2217-2242.
21. Statistical Center of Iran. Index of publications. [update 2008 October]. Available at URL: http://amar.sci.org.ir/
22. I.R of Iran Meteorological Organization (IRIMO). [update 2008 October]. Available at URL:http://www.irimo.ir/farsi/statistics/IT-Report/index.asp
23. Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C. Use of generalized linear mixed model in the spatial analysis of small area malaria incidence rate in KwaZulu Natal, South Africa. ***Am J Epidemiol*** 2001; 12: 1213-12121.
24. Szklo M, Nieto FJ. Epidemiology: beyond the basic. 2nd ed. Maryland (USA): Aspen Publishers Inc; 2007. p. 272-273.
25. McCulloch E, Searle SR. Generalized, Linear and Mixed Models. New York (NY): John Willy & Sons Inc; 2004.
26. Bruce-Chwatt LJ, editor. Essential malariology. 2nd ed. London (UK): William Heinemann Medical Books; 1985. p. 493.
27. Gosoniu L, Vounatsou P, Sogoba N, Smith T. Bayesian modeling of geo-statistical malaria risk data. ***Geospat Health*** 2006; 1: 127-139.

28. Haghdoost AA. Assessment of seasonal and climatic effects on the incidence and species composition of malaria by using GIS methods [Thesis]. London (UK): London School of Hygiene and Tropical Medicine; 2004. p. 85-97.
29. Haghdoost AA, Mazharia S, Bahaadinib K. Estimating the relapse risk of Plasmodium vivax in Iran under national chemotherapy scheme using a novel method. ***J Vector Borne Dis*** 2006; 43: 168-172.
30. Lind J. Essay on the diseases incidental to Europeans in hot countries, with the method of preventing their fatal consequences. 5th ed. London (UK): J Murray Publishing; 1792.
31. Cohen JM, Ernst KC, Lindblade KA, Vulule JM, John CC, Wilson ML. Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands. ***Malar J*** 2008; 7: 40.
32. Hay SI, Noor AM, Simba M, Busolo M, Guyatt HL, Ochola SA, et al. Clinical epidemiology of malaria in the highlands of western Kenya. ***Emerg Infect Dis*** 2002; 8: 543-548.
33. Thomson MC, Mason SJ, Phindela T, Connor SJ. Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. ***Am J Trop Med Hyg*** 2005; 73: 214-221.
34. Zhou G, Minakawa N, Githeko AK, Yan G. Association between climate variability and malaria epidemics in the East African highlands. ***Proc Natl Acad Sci USA*** 2004; 101: 2375-2380.
35. Nobre AA, Schmidt AM, Lopes HF. Spatio-temporal models for the mapping incidence of malaria in Para. ***Environmetrics*** 2005; 16: 291-304.
36. Mabaso MLH, Vounatsou P, Midzi S, Da Silva J, Smith T. Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe. ***Int J Health Geogr*** 2006; 5: 20.
37. Gomez-Elipe A, Otero A, Van-Herp M, Aguirre-Jaime A. Forecasting malaria incidence based on monthly case reports and environmental factors in Karuzi, Burundi, 1997-2003. ***Malar J*** 2007; 6: 129.

Related topics

Al-Tawfiq JA. Epidemiology of travel-related malaria in a non-malarious areas in Saudi Arabia. ***Saudi Med J*** 2006; 27: 86-89.

Alkhalife IS. Imported malaria infections diagnosed at the Malaria Referral Laboratory in Riyadh, Saudi Arabia. ***Saudi Med J*** 2003; 24: 1068-1072.

Alkhunaizi AM, Al-Tawfiq JA, Al-Shawaf MH. Transfusion-transmitted malaria in a kidney transplant recipient. How safe is our blood transfusion? ***Saudi Med J*** 2008; 29: 293-295.

Appendix A

Spatial Kriging: Assume suppose $Z(s_1), Z(s_2), \ldots, Z(s_n)$ are the outcome data associated with spatial locations $s_1, s_2, \ldots, s_n$. We are interested in predicting an unobserved variable $Z(s_0)$. An optimal predictor of $Z(s_0)$ can be written as

$$\hat{Z}(s_0) = \sum_{i=1}^{n} w_i Z(s_i)$$

Where the weights w_i are chosen to minimize the mean-squared prediction error. We chose a second order stationary process with marginal moments

$$E\big(Z(s)\big) \equiv m(s) = x'(s)\beta$$

$$cov\Big(Z(s_i), Z(s_j)\Big) = \sigma_\epsilon^2 + \sigma^2 \rho(s_i - s_j)$$

Where $x(s_i)$ is a $p \times 1$ vector of explanatory variables, β is a $p \times 1$ vector of unknown model parameters, $\sigma_\epsilon^2 + \sigma^2$ is the variance of the data, σ_ϵ^2 is a residual or nugget variance and $\rho(s_i - s_j)$ is the correlation function of the process.

GLM for spatial data: Linear models and linear predictors may not be the best choices for many applications, particularly when the data are binary or counts. In this context, GLMs are the most common alternative. We assume that some function of marginal mean, called the link function, related to the spatial surface, so that

$$g[m(s)] = x'(s)\beta$$

The two most common link functions are log for count data and the logit for binomial data. Since the SIR data has a count nature, the log transformation was used for modeling the malaria SIRs in this paper.

Thus, we used the following model for describing the statistical relationship between environmental explanatory variables and SIR values:

$$\log(z) = log\ (n) + x'(s)\boldsymbol{\beta} + \varphi_i$$

Where $\log(z)$ is the logarithm of the number of observed malaria cases, $log\ (n)$ is an offset, $\boldsymbol{\beta}$ consist of $\boldsymbol{\beta_0}$ as an intercept and $\boldsymbol{\beta_i}$ as regression coefficient vector corresponding to the environmental variables (precipitation, humidity, elevation and temperature), and φ_i is the spatial random effect for district i. In particular, if d_{ij} denotes the distance between points i and j, where observations z_i and z_j were made $\mathbf{V} = \mathbf{I}\sigma_1^2 + \mathbf{F}\sigma^2$, and $F_{ij} = \exp\ (-\frac{d_{ij}}{\rho})$. The unknown parameters in this model, σ_1^2 (the nugget), σ^2 (the sill) and ρ (the range), can be estimated from the variogram of the data as a vector of $\theta = \begin{bmatrix} \sigma_1^2 \\ \sigma^2 \\ \rho \end{bmatrix}$. For non-normally distributed data $\boldsymbol{\theta}$ is specified in the macro call to enable GLIMMIX to estimate the appropriate correlation matrix for spatially correlated Poisson model.

Appendix B

```
Attach (SBP)
library (gstat)
sir.var<- variogram(SIR~1, loc=~X+Y, SBP)
logsir.var<- variogram(log(SIR+0.000001)~1, loc=~X+Y, SBP)
print (sir.var)
print (logsir.var)
par (mfrow=c(2,2))
plot (sir.var)
plot (logsir.var)

logsircr.var<- variogram(log(SIR+0.000001)~1, loc=~X+Y,  SBP)
print (logsircr.var)
plot logsircr.var)
plot (logsircr.var$dist, logsircr.var $gamma, xlab="distance", ylab="Varigram")
plot (variogram(log(SIR+0.000001)~1, loc=~ X+Y, data=SBP, cloud=TRUE))

model.1 <- fit.variogram(logsircr.var, vgm(psill=1,model="Sph", range=100,nugget=1))
plot(logsircr.var, model=model.1)
title("Best Fit Spherical")

model.2 <- fit.variogram(logsircr.var, vgm(psill=12,model="Gau", range=45000,nugget=1))
model.3 <- fit.variogram(logsircr.var, vgm(psill=5,model="Exp", range=20000,nugget=1))
model.4 <- fit.variogram(logsircr.var, vgm(psill=10,model="Mat", range=20000,nugget=1))

m <- vgm(psill=130.02,"Exp", range=1324948,nugget=3.85)
logsircr.krige <- krige(log(SIR+0.000001)~1, ~X+Y, model = m, data = SBP,
newd = loc)
logsircr.krige
```