

A Bayesian-based approach for spatio-temporal modeling of county level prevalence of *Schistosoma japonicum* infection in Jiangsu province, China

Guo-Jing Yang^{a,b}, Penelope Vounatsou^b, Xiao-Nong Zhou^c, Marcel Tanner^b, Jürg Utzinger^{b,*}

^aJiangsu Institute of Parasitic Diseases, Wuxi, Jiangsu, Meiyuan 214064, People's Republic of China

^bDepartment of Public Health and Epidemiology, Swiss Tropical Institute, P.O. Box, CH-4002, Basel, Switzerland

^cNational Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai 200025, People's Republic of China

Received 24 August 2004; received in revised form 27 October 2004; accepted 2 November 2004

Abstract

Spatio-temporal variations of *Schistosoma japonicum* infection risk in Jiangsu province, China, were examined and the relationships between key climatic factors and infection prevalence at the county level were determined. The parasitological data were collected annually by means of cross-sectional surveys carried out in 47 counties from 1990 to 1998. Climatic factors, namely land surface temperature (LST) and normalized difference vegetation index (NDVI), were obtained from remote sensing satellite sensors. Bayesian spatio-temporal models were employed to analyze the data. The best fitting model showed that spatial autocorrelation in Jiangsu province decreased dramatically from 1990 to 1992 and increased gradually thereafter. A likely explanation of this finding arises from the large-scale administration of praziquantel for morbidity control of schistosomiasis. Our analysis suggested a negative association between NDVI and risk of *S. japonicum* infection. On the other hand, an increase in LST contributed to a significant increase in *S. japonicum* infection prevalence. We conclude that combining geographic information system, remote sensing and Bayesian-based statistical approaches facilitate integrated risk modeling of *S. japonicum*, which in turn is of relevance for allocation of scarce resources for control of schistosomiasis japonica in Jiangsu province and elsewhere in China, where the disease remains of public health and economic significance.

© 2004 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: Bayesian statistics; Conditional autoregressive model; Geographic information system; Remote sensing; *Schistosoma japonicum*; China

1. Introduction

It is widely acknowledged that schistosomiasis due to *Schistosoma japonicum* has been endemic in China for at least two millennia (Chen and Feng, 1999; Ross et al., 2001). Some 50 years ago, an estimated 12 million people were infected with *S. japonicum*, another 100 million people were at risk of acquiring the disease, and 14,000 km² of flood plains in the Yangtze River basin were identified as transmission zones (Chen and Feng, 1999; Ross et al., 2001). At that time, schistosome-attributable morbidity and mortality were high (Chen and Feng, 1999). In view of the great public health and economic impact the disease placed on the country, government officials embarked on vigorous research and disease control programmes.

A national sampling survey in 1995 revealed that the number of infected people was dramatically reduced to less than 1 million (Chen and Feng, 1999; Ross et al., 2001). The national schistosomiasis control programme of China, therefore, is considered to be one of the most successful ones in the world (Chitsulo et al., 2000; Engels et al., 2002; Utzinger et al., 2003). Despite substantial progress made over the past decades, transmission of *S. japonicum* still persists in seven provinces of China; namely, Anhui, Hubei, Hunan, Jiangsu and Jiangxi in the lake region, and Sichuan and Yunnan in the mountainous region. There is considerable concern that with the recent termination of a World Bank loan designated for schistosomiasis control, and with major water resource developments currently underway (i.e. Three Gorges dam and South-North water transfer project), the disease could re-emerge (Xu et al., 1999, 2000; Ross et al., 2001; Zheng et al., 2002).

* Corresponding author. Tel.: +41 61 284 8129; fax: +41 61 284 8105.
E-mail address: juerg.utzinger@unibas.ch (J. Utzinger).

Over the past decade, application of geographic information systems (GIS) and satellite remote sensing (RS) has become an integral part of China's national schistosomiasis control programme (Zhou et al., 2001; Yang et al., 2005). For example, areas at risk of high transmission intensities were identified, employing either the normalized difference vegetation index (NDVI), a climate risk index, or earth surface temperature and mean minimum temperature in January (Zhou et al., 1999). A comprehensive GIS database was established for the provinces of Anhui, Jiangsu and Jiangxi, and the annual *S. japonicum* infection prevalence data were analyzed longitudinally (Yang et al., 2002). Finally, the Chinese government now presents their annual reports on schistosome-endemic areas by means of digital maps, which aids the planning and operating of the national control programme (Chen et al., 2003).

It is important to note that most of the preceding work was based on collating, mapping and analyzing prevalence data, employing conventional statistical approaches. Only few attempts have been made to investigate spatio-temporal correlations and to assess how climatic factors affect these correlations. Furthermore, in areas with small population sizes, mapping of raw prevalence data can be confusing or even misleading, due to variability in the estimated prevalence. Consequently, it is difficult to tease apart chance variation from true differences (Sun et al., 2000). Bayesian spatio-temporal modeling has been recognized as a powerful means to provide more robust estimates, as it takes into account possible correlations and covariate effects, and allows to readily incorporating errors that may arise from mean or median estimates of the independent variables (Bernardinelli and Montomoli, 1992; Kleinschmidt et al., 2002; Basáñez et al., 2004; Dorny et al., 2004).

The purpose of this study is to investigate the relationship between key climatic factors derived from satellite RS and prevalence of *S. japonicum* infection obtained from repeated cross-sectional epidemiological surveys. The focus is on the Jiangsu province, located at the lower reaches of the Yangtze River. By the end of 2000, 35 counties reached the criterion of transmission interruption for *S. japonicum*, in eight counties transmission control was achieved, but in the remaining 14 counties, transmission still occurred (Chen and Feng, 1999). We employed annual county-level prevalence data for the period 1990–1998, and modeled the effects of two key climatic factors, utilizing a Bayesian approach to make predictions of infection risks over time.

2. Materials and methods

2.1. Approach and study area

In the study presented here, we combined GIS, satellite RS and Bayesian statistical methods for integrated spatio-temporal risk modelling of *S. japonicum*. The analysis is at

the county level, and the geographical focus is on Jiangsu province in eastern China.

2.2. Parasitological data

The prevalence data of *S. japonicum* infection were collected during cross-sectional, community-based surveys carried out annually between 1990 and 1998. At the beginning of the study, Jiangsu province contained 47 schistosome-endemic counties. The Jiangsu Institute of Parasitic Diseases (JIPD; Meiyuan, China) collected the data, adhering to standardized, quality-controlled field and laboratory procedures. In brief, all study participants initially underwent immunological examination. In a first step, a skin test was performed by injecting 0.03 ml of a schistosome egg or adult worm antigen. Those with a positive reaction were subsequently examined with the circum-oval precipitin test. Finally, when this second immunological test was positive, examination of a stool sample was carried out, employing the miracidium hatching test. Details of the immunodiagnostic techniques have been described elsewhere (Sleigh et al., 1998; Wu, 2002). According to Mao (1990), the miracidium hatching test is more sensitive than the Kato–Katz technique for diagnosis of *S. japonicum*, thus the former was utilized. Data management was also done by JIPD, and the database was kindly provided for the present study.

2.3. Climatic data

The climatic data utilized for the current study were (i) NDVI and (ii) land surface temperature (LST). They were derived from the Advanced Very High Resolution Radiometer (AVHRR), which is part of a polar-orbiting meteorological satellite. Ten-day composite images were downloaded from the AVHRR website, hosted by the US Geological Survey (USGS), for the period 1990–1998 (<http://edcdaac.usgs.gov/1KM/comp10d.asp>). These images were georeferenced and subsetting in ENVI version 3.5 software (Research System Inc.; Boulder, CO, USA). ArcGIS software (ESRI; Redlands, CA, USA) was used to extract average NDVI, band 4 (T4) and band 5 (T5) values, for each pixel of the image. The average was taken over the decadal values during the *S. japonicum* transmission season which occurs from April to October, separately for each year. The county-specific values were estimated as the median of all pixel values within the county.

The LST data were calculated for each year from the county-specific NDVI, T4 and T5 values, using the formula below, which was cited by Becker and Li (1990)

$$\begin{aligned} \text{LST} = & 1.274 + (T_4 + T_5)/2\{1 + [0.15616(1 - \varepsilon)/\varepsilon] \\ & - 0.482(\Delta\varepsilon/\varepsilon^2)\} + (T_4 - T_5)/2\{6.26 \\ & + [3.98(1 - \varepsilon)/\varepsilon] + 38.33(\Delta\varepsilon/\varepsilon^2)\} \end{aligned}$$

where $\Delta\varepsilon = 0.01019 + 0.01344 \ln(\text{NDVI})$, and $\varepsilon = 0.984605 + 0.0228 \ln(\text{NDVI})$.

2.4. Statistical analysis

The raw data were entered and validated, using internal consistency checks and analyzed in STATA version 8.0 (Stata Corporation; College Station, TX, USA) and WinBUGS version 1.4 (Imperial College and MRC, London, UK). In a first step, logistic regression analyses were performed to assess the bivariate relations between the infection prevalence of *S. japonicum* and the climatic covariates (Model 1). Next, a Bayesian spatio-temporal logistic model was fitted in WinBUGS (Model 2). Spatial correlation was modeled by county-specific random effects, which were assumed to arise from a conditional autoregressive model, CAR (γ), where γ measures the amount of the spatial correlation. Temporal correlation was modeled by yearly random effects, under the assumption that they follow an auto-regressive process, AR (1), with lag 1. Model 2 assumed independence between the spatial and temporal effects. Finally, Model 3 was applied, assuming a separate set of spatial random effects for each year. Bayesian smoothing was carried out by employing conditional autoregressive models, CAR (γ_t), for the random effects corresponding to year t . The deviance information criterion (DIC) (Spiegelhalter et al., 2002) was used to compare the goodness of fit of Models 2 and 3, and hence to select the model which better

quantifies the spatio-temporal correlation. The parameters of the models were estimated using Markov chain Monte Carlo (MCMC) simulation. Further details of the modeling approach are given in the Appendix.

3. Results

Table 1 shows parameter estimates from the non-spatial (Model 1), as well as the two spatio-temporal models (Models 2 and 3). The results obtained from the non-spatial analysis indicate that both climatic covariates investigated (i.e. LST and NDVI) are significantly related to the risk of *S. japonicum*. In particular, the infection prevalence of *S. japonicum* increases with decreasing values of NDVI (coef = -0.014 ; 95% confidence interval: -0.015 , -0.013), and with increasing values of LST (coef = 0.047 ; 95% confidence interval: 0.041 , 0.053).

Because this non-spatial model does not allow appraisal of possible spatio-temporal variation and accurate estimation of the standard errors of the parameters in the presence of these variations, spatio-temporal models were fitted to capture the space and time correlation in the data. Model 2 assumed that spatial correlation between counties remained unchanged during the entire study period.

Table 1

Estimates of different parameters for modeling country level *Schistosoma japonicum* prevalence data in Jiangsu province, China, by means of a non-spatial model and two different spatio-temporal models

Parameter	Model specifications		
	Non-spatial	Bayesian spatio-temporal	
	Model 1 ^a	Model 2	Model 3
Constant (α_0)	-17.28 (-19.16 , -15.40)	-5.43 (-6.79 , -4.74)	-1.69 (-3.21 , -0.55)
NDVI (α_1)	-0.014 (-0.015 , -0.013)	0.11 (0.085 , 0.12)	-0.093 (-0.097 , -0.085)
LST (α_2)	0.047 (0.041 , 0.053)	-0.069 (-0.077 , -0.052)	0.037 (0.033 , 0.040)
Temporal correlation (ρ)		-0.022 (-0.73 , 0.80)	
Temporal variation (σ^2)		0.76 (0.43 , 1.32)	
Spatial variation in 1990 (σ_1^2)		3.75 (2.94 , 4.80)	3.52 (2.66 , 4.68)
Spatial variation in 1991 (σ_2^2)			3.09 (2.42 , 3.97)
Spatial variation in 1992 (σ_3^2)			3.57 (2.79 , 4.62)
Spatial variation in 1993 (σ_4^2)			4.29 (3.34 , 5.56)
Spatial variation in 1994 (σ_5^2)			4.42 (3.31 , 5.93)
Spatial variation in 1995 (σ_6^2)			4.04 (3.00 , 5.46)
Spatial variation in 1996 (σ_7^2)			4.12 (3.07 , 5.57)
Spatial variation in 1997 (σ_8^2)			4.41 (3.19 , 6.11)
Spatial variation in 1998 (σ_9^2)			4.39 (3.13 , 6.22)
Spatial correlation in 1990 (γ_1)		0.89 (0.61 , 0.99)	0.89 (0.61 , 0.99)
Spatial correlation in 1991 (γ_2)			0.69 (0.10 , 0.97)
Spatial correlation in 1992 (γ_3)			5.87×10^{-4} (-0.85 , 0.77)
Spatial correlation in 1993 (γ_4)			0.20 (-0.64 , 0.84)
Spatial correlation in 1994 (γ_5)			0.77 (0.35 , 0.98)
Spatial correlation in 1995 (γ_6)			0.79 (0.37 , 0.98)
Spatial correlation in 1996 (γ_7)			0.91 (0.69 , 0.99)
Spatial correlation in 1997 (γ_8)			0.92 (0.75 , 0.99)
Spatial correlation in 1998 (γ_9)			0.91 (0.72 , 0.99)
Deviance information criterion (DIC)	$106,269.0^b$	$32,306.2$	2400.1

Brackets indicate 95% confidence intervals for Model 1 and 95% credible intervals for Model 2 and 3.

^a Parameters were estimated using the maximum likelihood method.

^b The DIC was obtained by fitting the Bayesian equivalent.

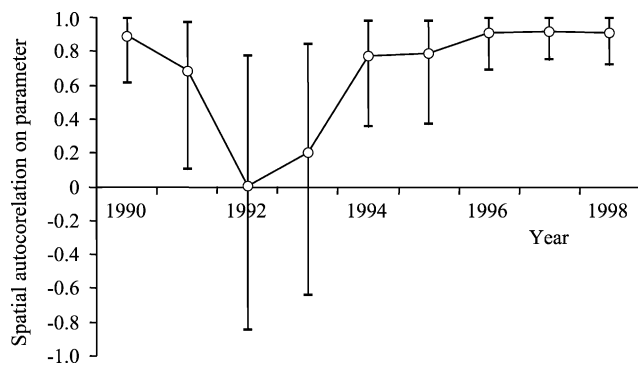


Fig. 1. Spatial autocorrelation parameter (γ_t) for county level prevalence of *Schistosoma japonicum* infection in Jiangsu province, China, from 1990 to 1998.

In addition, Model 2 included a temporal effect, which assumed that the risk of *S. japonicum* in a given year was only related to the one of the preceding year. In contrast, Model 3 assumed that the amount of spatial correlation was year-dependent. The DIC goodness of fit criterion suggested that Model 3 was superior to Model 2, as it resulted in a much smaller deviation between the observed and the fitted data ($\text{DIC}_{\text{Model 2}} = 32,306.2$ versus $\text{DIC}_{\text{Model 3}} = 2400.1$). Consequently, further interpretation of the results focused on Model 3.

Model 3 supported the results of the non-spatial model (Model 1), as each of the two covariates investigated showed a statistically significant association to the risk of *S. japonicum*. As before, NDVI was negatively associated with infection prevalence of *S. japonicum* ($\text{coef} = -0.093$;

95% credible interval: $-0.097, -0.085$), whereas a positive association was found between LST and infection prevalence of *S. japonicum* ($\text{coef} = 0.037$; 95% credible interval: $0.033, 0.040$).

The parameters γ_1 – γ_9 of Model 3 quantified the amount of spatial correlation of the infection risk between the counties over the 9-year observation period. The estimates of these parameters, including 95% credible intervals, are graphically depicted in Fig. 1. Results showed that the spatial correlation dramatically dropped from 0.89 in 1990 to 0.0006 in 1992. After 1992, a rapid increase was observed. During the last three years of the study, spatial correlation became stable, and quite narrow 95% credible intervals were observed.

The smoothed annual *S. japonicum* infection prevalence data for the 47 counties from 1990 to 1998 are illustrated in Fig. 2. The infection risk decreased substantially from 1990 to 1998, most notably after 1992. In 1998, most of the counties had reached the criterion of schistosomiasis control, according to the definition of Chen and Feng (1999). The majority of these controlled areas are situated far away from the Yangtze River. In contrast, during the entire study period, the counties along the Yangtze River were at the highest risk of *S. japonicum*. Fig. 2 also shows that there was one county in the southern part of Jiangsu province, in which the infection risk increased from the year 1991 to 1992, and only decreased after 1993 to approach low levels towards the end of the investigated period.

Fig. 3 shows the maps of the residual spatial variation after adjusting for the effects of the climatic covariates for

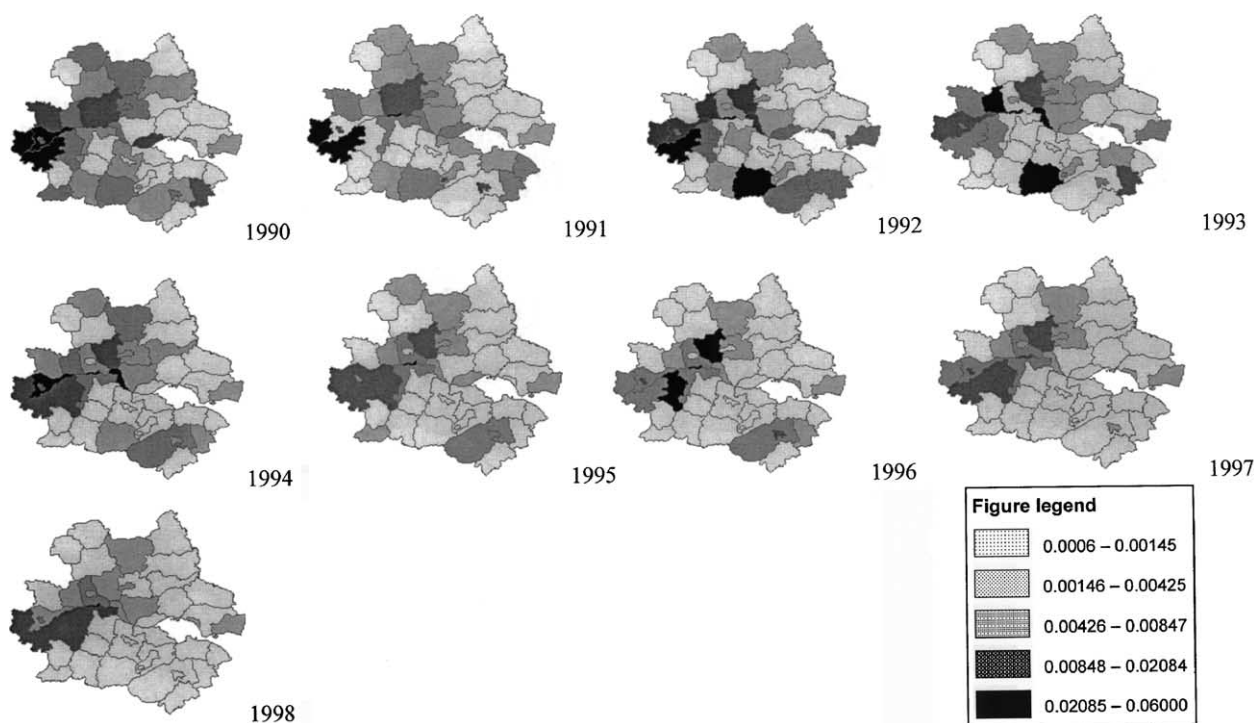


Fig. 2. Smoothed maps of prevalence of *Schistosoma japonicum* infection in Jiangsu province, China, from 1990 to 1998.

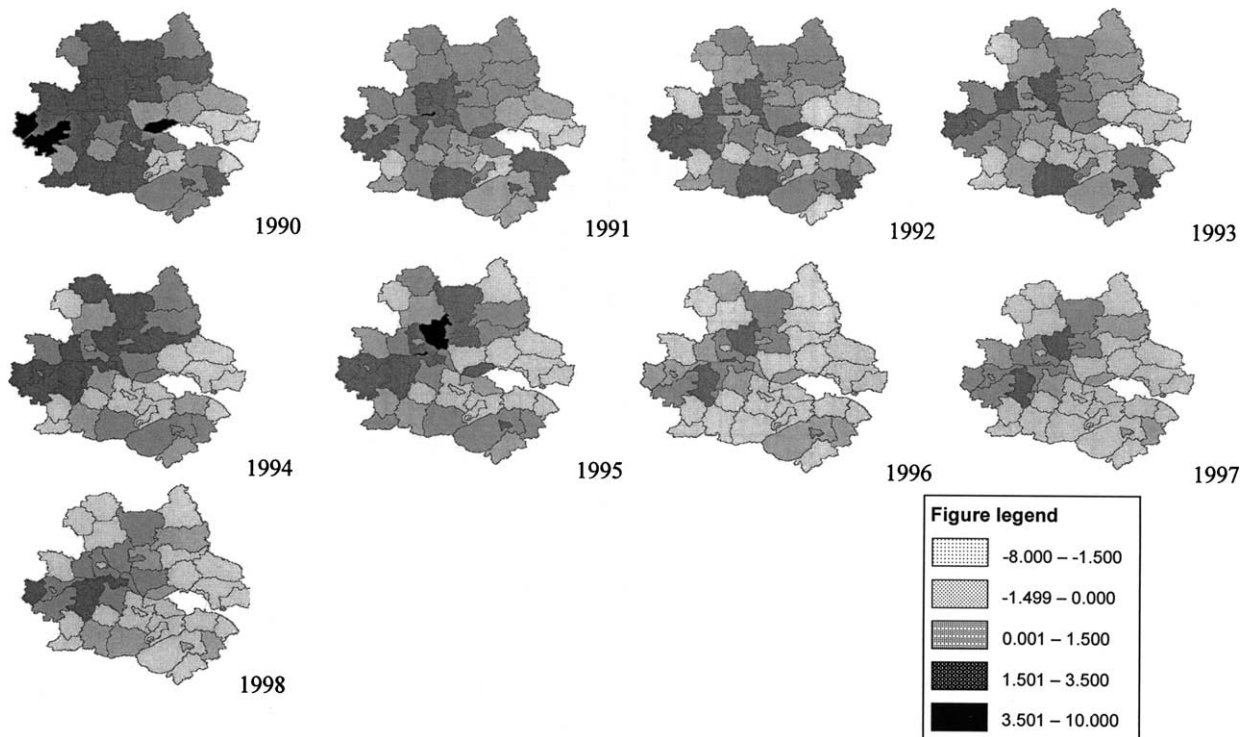


Fig. 3. Maps of the residual spatial variation after adjusting for the effects of the climatic covariates in Jiangsu province, China, from 1990 to 1998.

the study period of 1990 to 1998. It is apparent that in 1990, the counties with higher than average risk of *S. japonicum* are situated in the north-western part of the schistosomiasis-endemic area of the Jiangsu province. Since this pattern is less pronounced in Fig. 2, our modeling approach suggests that the climatic factors had a higher influence on that part of the province at the beginning of the study.

4. Discussion

Advances in GIS, RS and statistical methodology have opened new avenues to enhance our current understanding of spatio-temporal distribution of parasitic diseases (Hay, 2000; Malone et al., 2001; Zhou et al., 2001; Brooker et al., 2002). GIS and RS are convenient means for the creation of comprehensive geographical databases, and the display of the relevant data for policy discussion, guiding interventions and cost-effective monitoring at a range of spatial scales (Tatem and Hay, 2004). Sound statistical methods are mandatory to examine the underlying determinants of the observed patterns in the light of sampling variation and spatio-temporal correlations. Developments in Bayesian-based approaches and MCMC inference have greatly advanced spatio-temporal modeling (Gelfand and Smith, 1990). These methods have been successfully applied for parasitological diagnosis where there is a lack of a gold standard (Hubbard et al., 2002; Utzinger et al., 2002; Basáñez et al., 2004) and disease mapping (Cancré et al., 2000; Vounatsou et al., 2000; Assunção et al., 2001; Basáñez et al., 2004).

Schistosomiasis is a water-borne disease and its transmission is strongly correlated with environmental factors. We considered two of the most important factors, namely temperature and wetness, which characterize the overall climatic suitability for schistosomiasis transmission (Mao, 1990; Hay, 2000; Malone et al., 2001; Brooker et al., 2002). RS and GIS techniques were applied to extract the climatic factors (i.e. NDVI, LST) from satellite images (NOAA-AVHRR). NDVI maps indicate how much vegetation is present at each location with higher values representing more intensive vegetation coverage. Values range between -1 and $+1$, where the negative values correspond to water bodies. LST is one of the key parameters in the physics of land surface processes, combining surface-to-atmosphere interactions and the energy fluxes between the atmosphere and the ground. With a view to the biology and transmission dynamics of *S. japonicum* at the macro-scale, survival of intermediate host snails is compromised by decreasing temperatures. On the other hand, the availability of suitable water bodies is conducive for enhanced snail dispersal, which in turn can increase transmission intensity, as well as enlarge potential transmission areas. Therefore, LST is an important aggregate measure to predict both transmission areas and the intensity of transmission there, while NDVI is a key factor to predict the level of transmission. In summary, application of GIS and RS, by using NDVI and LST, is a viable approach for the assessment and prediction of *S. japonicum* infection prevalence at the county level. However, it is important to note that derived median estimates for the RS variables have some errors built around

them, which in turn can bias coefficients in the logistic regressions carried out.

The spatial resolution of the satellite images applied here was 8 km. In view of the inherent problem of scale in ecological studies (Levin, 1992), it is not practicable to predict the distribution of schistosomiasis at finer detail than at the county level. This is justified on three grounds. First, surface areas of the counties forming the schistosomendemic part of Jiangsu province range between 900 and 3000 km². Based on our spatial resolution, we were able to estimate climatic variables extracted from the satellite images with a pixel size of 64 km². Second, we employed the median for NDVI and LST estimates instead of the mean, because many water bodies in our study area revealed extreme values, which would have biased the results. Third, the information extracted from the satellite images was more accurate than the data provided by meteorological stations in terms of spatial resolution.

Both the non-spatial (Model 1) and the best fitting spatial-temporal model (Model 3) revealed a significant positive association between LST and county level *S. japonicum* infection prevalence, and a significant negative association between NDVI and *S. japonicum*. Interestingly, Model 2 also showed significant associations between LST and *S. japonicum*, and NDVI and *S. japonicum*, but with opposite signs. At first sight, we were surprised about this finding, particularly in view of the same set of covariates included in both Models 2 and 3. However, the two models capture the spatio-temporal correlation in a different way; hence our primary motivation to fit both models was to find out which model is superior in capturing this correlation. Clearly, Model 3 fits the data better than Model 2, as it resulted in a much smaller DIC. Therefore, it appears that there is no real temporal correlation but rather an important space–time interaction.

There is a plausible biological explanation for the significant positive association we found between LST and the risk of *S. japonicum*. The explanation arises from the parasite's life cycle, which consists of the stages of the egg, miracidium, sporocyst, cercaria, schistosomulum and the adult worm pair. Previous work has shown that the development of the parasite residing in the intermediate host snail is closely related to the environmental temperature. The higher the temperature, the more rapidly miracidia develop into cercariae, if thermal limits are not exceeded (Pesigan et al., 1958; Nagasaki, 1960). It follows that in areas with higher temperature, the parasite can complete its life cycle more promptly, and hence more cercariae are released into freshwater environments. As a result, the disease transmission intensity is enhanced, since cercariae are the infective stage for both humans and mammalian reservoir hosts.

The NDVI showed a statistically significant negative association with the infection prevalence of *S. japonicum*. This result means that areas with lower NDVI have higher chances for dispersal of intermediate host snails, thus

transmission intensity is increased and/or potential transmission areas are expanded. One possible explanation is that the intermediate host snail (i.e. *Oncomelania hupensis*) is amphibious. Areas where snail habitats are adjacent to water bodies are likely to be at an elevated risk of transmission, since low NDVI values indicate more water content. Hence, NDVI might be a good proxy for proximity to water bodies, which is a key factor in the transmission of *S. japonicum*. Another possible explanation is that areas with more dense vegetation tend to have lower LST and, as we argued above, lower temperatures were associated with a reduced risk of transmission. In fact, LST is usually considered as a viable climatic factor to delineate the critical limits beyond which disease transmission is not feasible any longer.

The results of the spatio-temporal analysis showed a decrease in the spatial autocorrelation during 1990 to 1992, reaching the lowest value for the whole study period in 1992. The smoothed *S. japonicum* prevalence map indicates that the endemic situation in 1992 was worse than in the previous year, which probably can be explained by a major flooding event of the Yangtze River in 1991. Recent work has shown that snail-infested areas enlarged and the risk of human–water contacts increased following flooding events in the preceding year (Zhou et al., 2002). From 1993 onwards, *S. japonicum* endemicity in the whole province became more stable, as indicated by the higher spatial correlation. The most likely explanation arises from reviewing the recent history of schistosomiasis control in China. With the overarching goal of eliminating schistosomiasis as a public health problem, a multi-year control programme was launched in 1992, financed through the World Bank and the Chinese government (Yuan et al., 2000). The programme was highly effective in reducing the public health significance of *S. japonicum* in China. However, control efforts in the schistosomendemic counties of the Jiangsu province along the Yangtze River proved particularly challenging. It is, therefore, not surprising that current estimates suggest that 95.7% of the snail-ridden areas and 85.6% of the *S. japonicum*-infected people are concentrated in the marshland and lake regions of the country (Yuan et al., 2002).

In view of ongoing ecological transformations as a result of major water resource developments (i.e. Three Gorges dam), it is conceivable that the marshland and lake regions will remain the 'hot spots' of *S. japonicum* transmission in China in the years to come. These activities can be further exacerbated by economic developments that are often paralleled by domestic migration of both people and cattle/water buffaloes that also serve as reservoir hosts of *S. japonicum* (Xu et al., 1999; Zheng et al., 2002). Meanwhile, another large-scale water resource development, namely the South–North water transfer project (<http://www1.people.com.cn/GB/shizheng/252/2283/>), and also global warming, are likely to further extend the schistosomendemic area further northwards. Hence, concerted efforts should be made to monitor and evaluate the effects of

these ecological transformations on the frequency and transmission dynamics of schistosomiasis japonica. Bayesian-based spatial statistical approaches hold promise for monitoring and evaluation purposes. Further studies are warranted to assess and quantify climatic and ecological factors other than LST and NDVI, as well as socioeconomic risk factors. Outcomes of such studies will be of relevance for more efficient and cost-effective resource allocation for control of schistosomiasis japonica in Jiangsu province and elsewhere in China, so that the public health and economic significance can be conquered.

Acknowledgements

This work received financial support from the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), project M8/181/4/Y.88 (ID-A10775), and the Chinese National Science Foundation (No. 300070684). Our sincere thanks go to the staff of the Department of Schistosomiasis Control, Jiangsu Institute of Parasitic Diseases, China. We thank Prof. Eberhard Parlow and Mr Gergely Rigo at the Institute for Meteorology, Climatology and Remote Sensing, University of Basel, for support on remote sensing issues, and two anonymous referees for a series of excellent comments. This work was carried out while P. Vounatsou was financially supported by the Swiss National Science Foundation (SNF; Project No. 3252B0-102136), and J. Utzinger in receipt of an ‘SNF-Förderungsprofessur’ (Project No. PPOOB-102883).

Appendix

Let Y_{it} and N_{it} denote the observed counts of schistosomiasis cases and survey population, respectively in county i and year t , and let P_{it} denote the prevalence of *Schistosoma japonicum* for the i th county in the t th year. We assumed that the Y_{it} 's are conditionally independent given, the N_{it} , P_{it} following a binominal distribution, i.e. $Y_{it} \sim \text{Bin}(N_{it}, P_{it})$ and introduced the covariate effects as well as the spatial and temporal random effects on the logit transformation of P_{it} . Model 1 assumes no spatial or temporal correlation, that is:

$$\log \text{it}(P_{it}) = \alpha_0 + \alpha_1 \times \text{NDVI}_{it} + \alpha_2 \times \text{LST}_{it} \quad (\text{Model 1})$$

where α_0 represents the mean prevalence over all counties and time periods, α_1 and α_2 are the coefficient of environmental factors NDVI and LST, respectively.

Model 2 includes spatial and temporal random effects, that is:

$$\log \text{it}(P_{it}) = \alpha_0 + \alpha_1 \times \text{NDVI}_{it} + \alpha_2 \times \text{LST}_{it} + \varphi_i + \omega_t \quad (\text{Model 2})$$

where φ_i is a random term that allows for spatially structured variation in the *S. japonicum* prevalence data and ω_t is a random term representing between-year variation. We introduce spatial dependence in the φ_i 's by assuming a conditional autoregressive model CAR (γ), which implies that each φ_i is conditional on the neighbour φ_j and follows a normal distribution with mean equal to the average of the neighbouring φ_j and variance equal to σ_i^2 scaled according to the number of the neighbours n_i of county i , that is $\varphi_i | \varphi_j, j \text{ neighbour of } i \sim N\left(1/n_i \gamma \sum_{j=1}^{n_i} \varphi_j, \sigma_i^2/n_i\right)$ where γ quantifies the strength of spatial dependence. We introduce the temporal dependence by assuming that ω_t 's following an autoregressive process with variance σ^2 where temporal correlation ρ exists only with the preceding year. Model 2 assumes that spatial correlation between the counties is independent of the year.

A separate model was fitted assuming that spatial correlations evolve over time, that is:

$$\log \text{it}(P_{it}) = \alpha_0 + \alpha_1 \times \text{NDVI} + \alpha_2 \times \text{LST} + \varphi_{it} \quad (\text{Model 3})$$

This Model 3 allows a different set of random spatial effects φ_{it} for each year t following a conditional autoregressive model $\varphi_{it} \sim \text{CAR}(\gamma_t)$ with spatial correlation parameter γ_t .

We fitted the above models using Bayesian statistical methods. According to the Bayesian approach we need to specify prior distributions for all model parameters. We adopted non-informative Uniform prior distributions for the covariate coefficients, i.e. vague inverse gamma priors for the variances σ^2 and $\sigma_i^2, i = 1, \dots, 9$ and vague normal priors for all other parameters.

References

- Assunção, R.M., Reis, I.A., Oliveira, C.D.L., 2001. Diffusion and prediction of leishmaniasis in a large metropolitan area in Brazil with a Bayesian space-time model. *Stat. Med.* 20, 2319–2335.
- Basáñez, M.-G., Marshall, C., Carabin, H., Gyorkos, T., Joseph, L., 2004. Bayesian statistics for parasitologists. *Trends Parasitol.* 20, 85–91.
- Becker, F., Li, Z.L., 1990. Towards a local split window method over land surfaces. *Int. J. Remote Sens.* 11, 369–393.
- Bernardinelli, L., Montomoli, C., 1992. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Stat. Med.* 11, 983–1007.
- Brooker, S., Beasley, M., Ndinarmotan, M., Madjiouroum, E.M., Baboguel, M., Djenguinabe, E., Hay, S.I., Bundy, D.A.P., 2002. Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bull. World Health Organ.* 80, 783–789.
- Cancré, N., Tall, A., Rogier, C., Faye, J., Sarr, O., Trape, J.-F., Spiegel, A., Bois, F., 2000. Bayesian analysis of an epidemiologic model of *Plasmodium falciparum* malaria infection in Ndiop, Senegal. *Am. J. Epidemiol.* 152, 760–770.
- Chen, M.G., Feng, Z., 1999. Schistosomiasis control in China. *Parasitol. Int.* 48, 11–19.

- Chen, X., Wu, X., Wang, L., Dang, H., Wang, Q., Zheng, J., Guo, J., Jiang, Q., Zhao, G., Zhou, X., 2003. Schistosomiasis situation in People's Republic of China in 2002. *Chin. J. Schisto. Control* 15, 241–244.
- Chitsulo, L., Engels, D., Montresor, A., Savioli, L., 2000. The global status of schistosomiasis and its control. *Acta Trop.* 77, 41–51.
- Dorny, P., Phiri, I.K., Vercruyse, J., Gabriel, S., Willingham III., A.L., Brandt, J., Victor, B., Speybroeck, N., Berkvens, D., 2004. A Bayesian approach for estimating values for prevalence and diagnostic test characteristics of porcine cysticercosis. *Int. J. Parasitol.* 34, 569–576.
- Engels, D., Chitsulo, L., Montresor, A., Savioli, L., 2002. The global epidemiological situation of schistosomiasis and new approaches to control and research. *Acta Trop.* 82, 139–146.
- Gelfand, A.E., Smith, A.F.M., 1990. Sampling based approaches to calculating marginal densities. *J. Am. Stat. Assoc.* 85, 398–409.
- Hay, S.I., 2000. An overview of remote sensing and geodesy for epidemiology and public health application. *Adv. Parasitol.* 47, 1–35.
- Hubbard, A., Liang, S., Maszle, D., Qiu, D., Gu, X., Spear, R.C., 2002. Estimating the distribution of worm burden and egg excretion of *Schistosoma japonicum* by risk group in Sichuan Province, China. *Parasitology* 125, 221–231.
- Kleinschmidt, I., Sharp, B., Mueller, I., Vounatsou, P., 2002. Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends. *Am. J. Epidemiol.* 155, 257–264.
- Levin, S.A., 1992. The problem of pattern and scale in ecology. *Ecology* 73, 1943–1967.
- Malone, J.B., Yilma, J.M., McCarroll, J.C., Erko, B., Mukaratirwa, S., Zhou, X.Y., 2001. Satellite climatology and the environmental risk of *Schistosoma mansoni* in Ethiopia and East Africa. *Acta Trop.* 79, 59–72.
- Mao, C.P., 1990. Biology of Schistosome and Control of Schistosomiasis. People's Health Press, Beijing.
- Nagasaki, M., 1960. Studies on the development of *Schistosoma japonicum* larva within the snail host, *Oncomelania nosophora*, with special reference to temperature. *Jap. J. Pub. Health* 15, 74–95.
- Pesigan, T.P., Hairston, N.G., Jauregui, J.J., Garcia, E.G., Santos, A.T., Santos, B.C., Besa, A.A., 1958. Studies on *Schistosoma japonicum* infection in the Philippines. 2. The molluscan host. *Bull. World Health Org.* 18, 481–578.
- Ross, A.G.P., Sleight, A.C., Li, Y.S., Davis, G.M., Williams, G.M., Jiang, Z., Feng, Z., McManus, D.P., 2001. Schistosomiasis in the People's Republic of China: prospects and challenges for the 21st century. *Clin. Microbiol. Rev.* 14, 270–295.
- Sleight, A., Li, X.M., Jackson, S., Huang, K.L., 1998. Eradication of schistosomiasis in Guangxi, China. Part 1: setting, strategies, operations, and outcomes, 1953–92. *Bull. World Health Org.* 76, 361–372.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* 64, 583–639.
- Sun, D.C., Tsutakawa, R.K., Kim, H., He, Z.Q., 2000. Spatio-temporal interaction with disease mapping. *Stat. Med.* 19, 2015–2035.
- Tatem, A.J., Hay, S.I., 2004. Measuring urbanization pattern and extent for malaria research: a review of remote sensing approaches. *J. Urban Health* 81, 363–376.
- Utzinger, J., Vounatsou, P., N'Goran, E.K., Tanner, M., Booth, M., 2002. Reduction in the prevalence and intensity of hookworm infections after praziquantel treatment for schistosomiasis infection. *Int. J. Parasitol.* 32, 759–765.
- Utzinger, J., Bergquist, R., Xiao, S.H., Singer, B.H., Tanner, M., 2003. Sustainable schistosomiasis control—the way forward. *Lancet* 362, 1932–1934.
- Vounatsou, P., Smith, T., Gelfand, A.E., 2000. Spatial modelling of multinomial data with latent structure; an application to geographical mapping of human gene and haplotype frequencies. *Biostatistics* 1, 177–189.
- Wu, G.L., 2002. A historical perspective on the immunodiagnosis of schistosomiasis in China. *Acta Trop.* 82, 193–198.
- Xu, X.J., Yang, X.X., Dai, Y.H., Yu, G.Y., Chen, L.Y., Su, Z.M., 1999. Impact of environmental change and schistosomiasis transmission in the middle reaches of the Yangtze River following the Three Gorges construction project. *Southeast Asian J. Trop. Med. Public Health* 30, 549–555.
- Xu, X.J., Wei, F.H., Yang, X.X., Dai, Y.H., Yu, G.Y., Chen, L.Y., Su, Z.M., 2000. Possible effects of the Three Gorges dam on the transmission of *Schistosoma japonicum* on the Jiang Han plain, China. *Ann. Trop. Med. Parasitol.* 94, 333–341.
- Yang, G.J., Zhou, X.N., Wang, T.P., Lin, D.D., Hu, F., Hong, Q.B., Sun, L.P., 2002. Establishment and analysis of GIS databases on schistosomiasis in three provinces in the lower reaches of the Yangtze River. *Chin. J. Schisto. Control* 14, 21–24.
- Yang, G.J., Vounatsou, P., Zhou, X.N., Utzinger, J., Tanner, M., 2005. A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in China. *Acta Trop.* 2005; (in press).
- Yuan, H.C., Guo, J.G., Bergquist, R., Tanner, M., Chen, X.Y., Wang, H.Z., 2000. The 1992–1999 World Bank schistosomiasis research initiative in China: outcome and perspectives. *Parasitol. Int.* 49, 195–207.
- Yuan, H., Jiang, Q., Zhao, G., He, N., 2002. Achievements of schistosomiasis control in China. *Mem. Inst. Oswaldo Cruz* 97 (Suppl. 1), 187–189.
- Zheng, J., Gu, X.G., Xu, Y.L., Ge, J.H., Yang, X.X., He, C.H., Tang, C., Cai, K.P., Jiang, Q.W., Liang, Y.S., Wang, T.P., Xu, X.J., Zhong, J.H., Yuan, H.C., Zhou, X.N., 2002. Relationship between the transmission of schistosomiasis japonica and the construction of the Three Gorge reservoir. *Acta Trop.* 82, 147–156.
- Zhou, X.N., Hu, X., Sun, N., Hong, Q., Sun, L., Lu, G., Fuentes, M., Malone, J.B., 1999. Application of geographic information systems on schistosomiasis surveillance. II. Predicting transmission intensity. *Chin. J. Schisto. Control* 11, 66–70.
- Zhou, X.N., Malone, J.B., Kristensen, T.K., Bergquist, N.R., 2001. Application of geographic information systems and remote sensing to schistosomiasis control in China. *Acta Trop.* 79, 97–106.
- Zhou, X.N., Lin, D.D., Yang, H.M., Chen, H.G., Sun, L.P., Yang, G.J., Hong, Q.B., Brown, L., Malone, J.B., 2002. Use of Landsat TM satellite surveillance data to measure the impact of the 1998 flood on snail intermediate host dispersal in the lower Yangtze River basin. *Acta Trop.* 82, 199–205.