

# Bayesian spatial modelling and the significance of agricultural land use to scrub typhus infection in Taiwan

Nicola A. Wardrop<sup>1</sup>, Chi-Chien Kuo<sup>1</sup>, Hsi-Chieh Wang<sup>2</sup>, Archie C. A. Clements<sup>3</sup>, Pei-Fen Lee<sup>4</sup>, Peter M. Atkinson<sup>1</sup>

<sup>1</sup>Geography and Environment, University of Southampton, Hampshire, UK; <sup>2</sup>Research and Diagnostic Centre, Centres for Disease Control, Department of Health, Taipei, Taiwan; <sup>3</sup>School of Population Health, University of Queensland, Brisbane, Australia; <sup>4</sup>Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan

**Abstract.** Scrub typhus is transmitted by the larval stage of trombiculid mites. Environmental factors, including land cover and land use, are known to influence breeding and survival of trombiculid mites and, thus, also the spatial heterogeneity of scrub typhus risk. Here, a spatially autoregressive modelling framework was applied to scrub typhus incidence data from Taiwan, covering the period 2003 to 2011, to provide increased understanding of the spatial pattern of scrub typhus risk and the environmental and socioeconomic factors contributing to this pattern. A clear spatial pattern in scrub typhus incidence was observed within Taiwan, and incidence was found to be significantly correlated with several land cover classes, temperature, elevation, normalized difference vegetation index, rainfall, population density, average income and the proportion of the population that work in agriculture. The final multivariate regression model included statistically significant correlations between scrub typhus incidence and average income (negatively correlated), the proportion of land that contained mosaics of cropland and vegetation (positively correlated) and elevation (positively correlated). These results highlight the importance of land cover on scrub typhus incidence: mosaics of cropland and vegetation represent a transitional land cover type which can provide favourable habitats for rodents and, therefore, trombiculid mites. In Taiwan, these transitional land cover areas tend to occur in less populated and mountainous areas, following the frontier establishment and subsequent partial abandonment of agricultural cultivation, due to demographic and socioeconomic changes. Future land use policy decision-making should ensure that potential public health outcomes, such as modified risk of scrub typhus, are considered.

**Keywords:** scrub typhus, spatial epidemiology, conditionally autoregressive model, disease ecology, *Orientia tsutsugamushi*, Taiwan.

## Introduction

Scrub typhus, caused by the rickettsia *Orientia tsutsugamushi*, is a lethal febrile human disease occurring throughout the western Pacific, south Asia and north-eastern Australia, with more than one billion people at risk of infection (Kawamura et al., 1995; Rosenberg, 1997). Scrub typhus is transmitted by the larval stage of trombiculid mites (i.e. chiggers) infected with *O. tsutsugamushi* and clinical presentation may begin with an eschar (ulcer) at the site of the infective chigger bite (this is not present in all cases), proceeding to non-specific symptoms such as headache, fever, muscle pains and cough (Seong et al.

2001). More severe symptoms and complications can occur in some patients, including haemorrhage, meningitis, pneumonia and multi-organ failure: it is thought that the severity of symptoms differs according to the strain of *O. tsutsugamushi* (Seong et al., 2001).

It is well recognised that the occurrence of many vector-borne diseases is typically associated with specific land cover or land use types (Patz et al., 2004; Lambin et al., 2010). This is hardly surprising given that disease vectors are primarily arthropods, with optimal development, survival and reproduction requiring particular physical environmental characteristics that determine specific land cover or land use types. For example, larvae of the mosquito *Anopheles darlingi*, the most important malaria vector in South America, are more abundant in deforested than forested areas in the Peruvian Amazon, likely due to the increased availability of breeding sites for *A. darlingi* following human settlement (Vittor et al., 2009). Moreover, wild vertebrate hosts, required by many disease vectors as sources of nour-

Corresponding author:  
Nicola A. Wardrop  
Geography and Environment  
University of Southampton, Highfield Campus, Building 44  
University Road, Southampton, Hampshire SO17 1BJ, UK  
Tel. +44 23 8059-4612; Fax +44 23 8059-3295  
E-mail: Nicola.Wardrop@soton.ac.uk

ishment or pathogen infection, normally flourish in particular habitats. As an example, density and infection prevalence of nymphs of the *Ixodes scapularis* tick, the primary vector for Lyme disease in northeast of the United States of America (USA), were higher in fragmented than continuous forests, which could be caused by the higher density of the white-footed mouse *Peromyscus leucopus* (the reservoir host) in smaller forest fragments (Allan et al., 2003). Identifying associations between land cover and risk of vector-borne diseases provides insight into the underlying processes of disease occurrence and could have profound significance for land use policy (Patz et al., 2004; Arinaminpathy et al., 2009; Lambin et al., 2010). For instance, infective rates of West Nile virus in competent *Culex* mosquitoes are lower in urban wetlands than adjacent residential areas; this is related to the wetlands harbouring lower proportions of competent avian and mosquito species and suggests that preserving urban wetlands may lower the human risk to West Nile virus infection (Johnson et al. 2012). Similarly, the higher risk to Lyme disease in forest fragments than in continuous forests in northeast USA (Allan et al., 2003) calls for a re-consideration of low-density residential land use, which not only furthers the severance of forests but also bring humans into close contact with infected ticks when dispersed houses are surrounded with fragmented forests (Jackson et al., 2006). However, detailed investigations for many vector-borne diseases remain unavailable (Patz et al., 2004). Studies are particularly urgently needed in Asia where rapid economic growth has dramatically transformed the land cover (Zhao et al., 2006).

The life cycle of trombiculid mites includes seven stages (egg, deutovum, larva (chigger), protonymph, deutonymph, tritonymph and adult), of which only chiggers are parasitic. During their nymph and adult stages they are free-living on the ground and feed on the eggs and larvae of arthropods, while they feed predominantly on rodents during their larval stage (Kawamura et al., 1995). Thus, their distribution is constrained by the availability of suitable hosts (i.e. rodents) along with environmental conditions that can affect their survival (Traub and Wisseman, 1974). Chigger abundance is limited by temperature (particularly in temperate areas) and precipitation (particularly in tropical areas) (Gentry et al., 1977). Land cover is also an important factor in determining the occurrence of the disease, with higher incidence occurring in areas of secondary or transitional vegetation, presumably in relation to the pre-

ferred habitat of the rodent species on which chiggers generally feed (Traub and Wisseman, 1974). In addition, socioeconomic factors may influence transmission: higher incidence tends to occur in rural areas, with farmers at higher risk of infection, and it has been shown that increasing urbanisation and school enrolment can decrease incidence of scrub typhus (Olson and Bourgeois, 1979; Ogawa et al., 2002; Lee et al., 2006; Vallee et al., 2010; Kuo et al., 2011). Thus, we can assume that the spatial distribution of scrub typhus is determined by climatic and environmental factors (including land cover and land use), which influence the distribution of the trombiculid mites and their preferred meal sources and socioeconomic factors, which may alter the pattern of contact between humans and chiggers.

Very few studies have analysed the spatial distribution of scrub typhus incidence in relation to climatic and environmental factors, particularly the influence of land cover/use, despite the importance of these factors for disease transmission and their potential implications for land use policy. Kuo et al. (2011) quantified the association of scrub typhus incidence with temperature, rainfall, and the normalized difference vegetation index (NDVI) (a measure of the greenness of vegetation derived from satellite remote sensing) in Taiwan but did not include land cover in the analyses. In Taiwan and Lao People's Democratic Republic, risk of scrub typhus infection was investigated only for a limited range of land use types (forest and degree of urbanization, respectively) (Vallee et al., 2010; Tsai and Yeh, 2013).

Here, we present a spatial analysis of scrub typhus incidence in Taiwan between 2003 and 2011 that considers land cover, climatic and socioeconomic variables in a spatially explicit analysis. Moreover, different from our previous study (Kuo et al., 2011), a Bayesian framework, with inclusion of spatially correlated and non-correlated random effects, was applied to more rigorously estimate spatial variation in disease risk, while quantifying covariate effects. This type of approach is commonly used to estimate spatial variation in disease risk within small areas (e.g. administrative areas) and ensures full representation of uncertainty in parameter estimates (Best et al., 2005; Lawson, 2009). This research, thus, explores the influence of land cover on the spatial distribution of scrub typhus transmission in Taiwan, and provides increased understanding of the spatial pattern of scrub typhus risk and the environmental and socioeconomic factors contributing to this pattern.

## Materials and methods

### *Study area*

The research focused on the main island of Taiwan, which has an area of 36,000 km<sup>2</sup>. The smaller associated islands were excluded due to ecological differences from the main island, which may result in differing epidemiological patterns. Taiwan experiences a marine tropical climate with hot, humid summers and cooler, drier winters. The western plains of Taiwan house the majority of the population, while the mountainous areas in central and eastern Taiwan have lower population densities. The smallest administrative areas to which scrub typhus diagnoses can be linked are administrative districts (in urban areas) and townships (in rural areas): these were the administrative areas used for the purposes of analysis.

### *Scrub typhus data*

Data regarding numbers of scrub typhus diagnoses between 2003 and 2011 were obtained from the Taiwan Centers for Disease Control: the disease is notifiable within Taiwan. Human infections were confirmed by using real-time polymerase chain reaction (PCR) or an indirect immunofluorescent assay (IFA). The scrub typhus records were linked to the administrative unit in which the infection was presumed to have been acquired (this information has been recorded routinely since 2003) instead of the address of residence, as infection is typically acquired away from the house (Kawamura et al., 1995), to produce a count of incident scrub typhus cases in each administrative area. Population data (2003 to 2011) were obtained from the Department of Statistics of the Taiwan Ministry of the Interior: these were also used to calculate population density within administrative areas (persons per km<sup>2</sup>).

### *Covariate data*

Climatic datasets were obtained from the Taiwan Central Weather Bureau: long-term averages (from 1959 to 1985) were used to provide minimum, mean and maximum temperature (°C) and total annual rainfall (mm) at a spatial resolution of 1 x 1 km (Lee et al., 1997). The NDVI was calculated using a mosaic of *Système pour L'Observation de la Terre* (SPOT) images from 2002 at a spatial resolution of 1 x 1 km. Elevation data (in m) were obtained from the Shuttle Radar Topography Mission (SRTM), with a spatial

resolution of 3 arc sec. Average values for the administrative areas were extracted from the temperature, rainfall, NDVI and elevation datasets using ArcMap 10.0 (ESRI; Redlands, USA). Land cover data were obtained from Globcover (European Space Agency, 2008) and the initial land cover classes were merged to create a small number of land cover types thought to be important for scrub typhus transmission: mosaics of forest, grass and shrubs (hereafter referred to as forest-grass-shrubs); croplands; mosaics of crops and vegetation (including grassland, shrubland and forest; referred to as crop-vegetation); and forest. The proportion of each administrative area that consisted of each of these land cover classes was calculated using ArcMap 10.0 (ESRI) to provide a quantitative measure of the land cover.

The proportion of the population within each administrative area that worked in agriculture was obtained from the Taiwan Council of Agriculture 1996 to 2004 (referred to as proportion in agriculture). The average income (thousands of New Taiwan dollar (NTD) per year) along with the median and standard deviation for the year 2005 for each administrative area were obtained from the Fiscal Information Agency of the Taiwan Ministry of Finance to provide a measure of socioeconomic status.

### *Exploratory analysis*

The correlation between scrub typhus human infection and each of the covariates listed above was examined using scatterplots and univariate Poisson regression in the R statistical package. Variables that were significantly correlated with the outcome were assessed for co-linearity with other significant variables. Where two (or more) variables that were measures of related phenomena (e.g. minimum temperature and average temperature) produced a large correlation (>0.75), the variable with the smallest value of the Akaike information criterion (AIC) in a univariate regression model was selected for use and the others omitted. Where two (or more) variables were strongly correlated, but were considered to be measures of unrelated phenomena, these covariates were used in model comparison in the non-spatial Poisson regression analysis (see below).

### *Non-spatial Poisson regression analysis*

Multivariate Poisson regression was carried out using the covariates found to be significantly associated with the outcome during the exploratory analysis

(P-value  $\leq 0.2$  from univariate Poisson regression). Candidate models contained one of the covariates marked as collinear to one another in Table 1, along with the all the other, non-collinear, covariates in Table 1. Any variable, which was not statistically significant ( $P > 0.05$ ) was subsequently removed from the model; if multiple variables were non-significant, the variable with the highest p-value was removed first, significance assessed for the remaining variables, and then further variables were removed in the same way if necessary. Models containing population density, elevation and maximum temperature as quadratic expressions were also included in the comparison; the quadratic expression was retained in the model if it resulted in an improved model fit (as shown by a decrease in AIC). To check for spatial autocorrelation in the residuals from the multivariate Poisson model (which would indicate that the covariates in the model had not fully captured the spatial pattern in scrub typhus human infection), Moran's  $I$  was calculated using the Pearson's residuals.

#### *Spatial Poisson regression analysis*

The influence of spatial autocorrelation (where observations close together in space are more likely to have similar values than observations that are far apart) was addressed, and overdispersion in the data accounted for, using a Poisson log-normal regression model with a non-spatial random effect and a spatially dependent random effect structured using a conditionally autoregressive (CAR) prior (Lawson, 2009; Kéry, 2010). The CAR component utilised a simple spatial adjacency weight matrix to represent spatial relationships between administrative areas. Non-spa-

tial random effects were also included in the model comparison. The covariates selected in the non-spatial multivariate Poisson regression analysis described above were included as fixed effects in this stage of the analysis.

The modelling, carried out in WinBUGS, assumed that the observed count of scrub typhus cases  $y_i$  for administrative area  $i$  ( $i = 1$  to 349) followed a Poisson distribution with a mean of  $\mu_i$ :

$$y_i \sim \text{Poisson}(\mu_i)$$

and

$$\log \mu_i = \log E_i + \theta_i$$

where  $E_i$  (the expected number of cases of scrub typhus within administrative area  $i$ ) is an offset term used to control for population size within the administrative areas and  $\theta_i$  the log relative risk. The log relative risk  $\theta_i$  was modelled as:

$$\theta_i = \alpha + \beta_{1\dots i} C_{1\dots i} + S_i + U_i$$

where  $\alpha$  is the intercept term;  $\beta_{1\dots i}$  a vector of coefficients for the covariates,  $C_{1\dots i}$ ;  $S_i$  a spatially correlated random effect structured using a CAR prior, which was assigned a simple spatial adjacencies matrix (to identify which polygons are directly adjacent to one another) and precision  $\tau$ ; and  $U_i$  a non-spatially correlated random effect with mean zero and precision  $\tau$ . A non-informative, flat prior was assigned to  $\alpha$ , while normally distributed priors (with mean zero and precision 0.001) were assigned to  $\beta_{1\dots i}$ . The  $\tau$  parameters for the spatial and non-spatial random effects were assigned non-informative gamma priors.

Table 1. Regression coefficients, rate ratios, 95% confidence intervals (CI) and other summary information from the univariate Poisson regression analysis.

| Variable  | Regression coefficient | Rate ratio (95% CI) | P-value | AIC <sup>a</sup> |
|---|------------------------|---------------------|---------|------------------|
| Maximum temperature* (°C)                                     | -0.44                  | 0.64 (0.64-0.65)    | <0.005  | 5,718.5          |
| Elevation* (100 m increase)                                   | 0.23                   | 1.26 (1.25-1.27)    | <0.005  | 5,779.4          |
| NDVI* (0.01 increase)   | 0.12                   | 1.12 (1.12-1.13)    | <0.005  | 5,538.4          |
| Rainfall (100 mm increase)                                    | 0.05                   | 1.05 (1.05-1.06)    | <0.005  | 7,981.5          |
| Proportion in agriculture (10% increase)                      | 0.36                   | 1.44 (1.41-1.46)    | <0.005  | 6,710.6          |
| Population density (100 persons per km <sup>2</sup> increase) | -0.01                  | 0.99 (0.98-0.99)    | <0.005  | 7,190.9          |
| Crop-vegetation (0.01% increase)                              | 0.09                   | 1.09 (1.09-1.09)    | <0.005  | 6,456.7          |
| Forest* (0.1% increase)                                       | 0.30                   | 1.35 (1.33-1.37)    | <0.005  | 6,553.4          |
| Forest-grass-shrub (1% increase)                              | 0.16                   | 1.17 (0.74-1.82)    | 0.490   | 8,267.2          |
| Croplands (1% increase)                                       | -1.09                  | 0.34 (0.28-0.40)    | <0.005  | 8,090.2          |
| Average income (100 NTD per year increase)                    | -0.77                  | 0.46 (0.44-0.48)    | <0.005  | 6,634.8          |

<sup>a</sup>Akaike information criterion; \*denotes covariates that are collinear to one another. Note also that the scale differs between variables.

Following an initial burn-in period of 10,000 iterations, a further 200,000 iterations were performed, and every tenth iteration was stored for parameter estimation to reduce autocorrelation in the samples. Markov chain Monte Carlo (MCMC) chains were inspected to check for convergence and multiple initial values for each parameter were used to ensure that the MCMC algorithms were converging on the same parameter space from different starting points. Summary measures for the posterior distribution of each parameter (posterior mean, standard deviation and 95% credible interval [CrI]) were stored to provide parameter estimates.

Four candidate models were compared: (i) fixed effects only; (ii) fixed effects and non-spatial random effects; (iii) fixed effects and spatially correlated random effects; and (iv) fixed effects, spatially correlated random effects and non-spatial random effects. Models were compared using the deviance information criterion (DIC): the model with the lowest DIC model was selected as the final, best fitting, model. All covariates, including those which did not maintain statistical significance ( $P > 0.05$ ), were retained to provide a summary of all covariate effects.

## Results

### Exploratory analysis

From 2003 to 2011, a total of 2,196 confirmed scrub typhus infections were reported on the main island of Taiwan, with an overall mean incidence rate of 1.07 cases per year per 100,000 residents. Within the 349 administrative areas of Taiwan's main island, cumulative case counts ranged from 0 to 106, with an average of 6.3 and a median of 2.0 (incidence rate per 100,000 person-years ranged from 0 to 75.9, with an average of 4.2 and median of 0.45). A spatial pattern was apparent in the annual incidence rate across the study area (Fig. 1), with higher incidence in the larger, more rural administrative areas in the east and central parts of the island. A comparison of Figs. 1 and 2 indicates high scrub typhus incidence tended to occur in areas with more forest cover: these areas are also cooler, at higher elevation, and have higher average NDVI values.

All the covariates listed in Table 1 were found to be significantly correlated with scrub typhus using univariate non-spatial Poisson regression ( $P < 0.05$ ), with the exception of forest-grass-shrubs ( $P = 0.490$ ). The

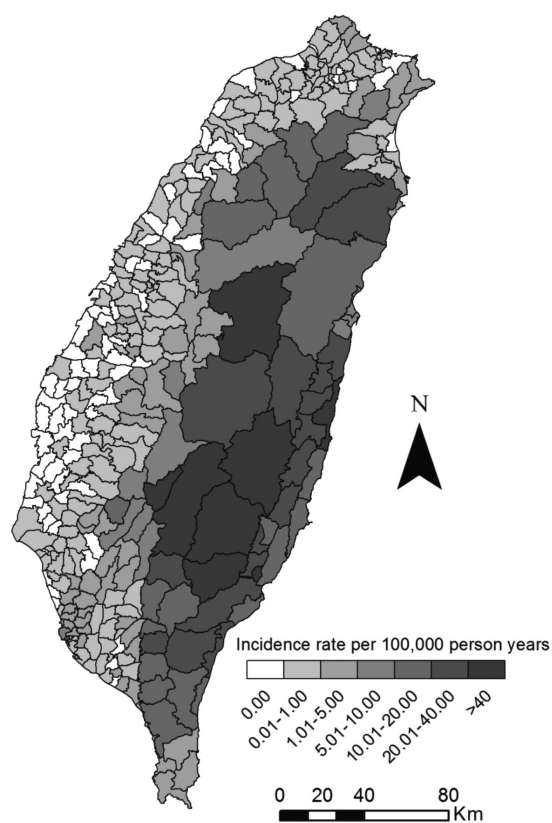


Fig. 1. Incidence rate per 100,000 person-years of scrub typhus within administrative areas in Taiwan, 2003-2011.

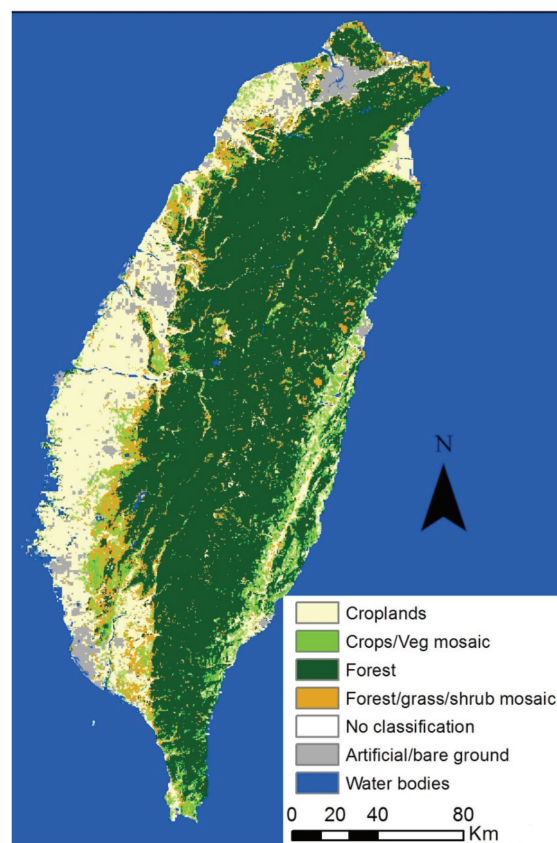


Fig. 2. Land cover in Taiwan (classes derived from Globcover data)

Table 2. Final multivariate Poisson regression model with coefficients, rate ratios, 95% confidence intervals and P-values.

| Variable                                   | Regression coefficient | Rate ratio (95% CI) | P-value |
|--|------------------------|---------------------|---------|
| Intercept                                  | -8.47                  |                     |         |
| Proportion in agriculture (10% increase)   | -0.12                  | 0.88 (0.85-0.91)    | < 0.005 |
| Crop-vegetation (0.01% increase)           | 0.09                   | 1.09 (1.09– 1.10)   | < 0.005 |
| Income-average (100 NTD per year increase) | -0.25                  | 0.78 (0.74-0.82)    | < 0.005 |
| Elevation (100 m increase)                 | 0.39                   | 1.48 (1.43-1.53)    | < 0.005 |
| Elevation*                                 | -0.01                  | 0.99 (0.99-0.99)    | < 0.005 |

\* Elevation was included as a quadratic expression in this model. Note that the scale differs between variables.

covariates minimum temperature, mean temperature, median income and standard deviation of income were excluded due to collinearity with other, more significant variables that were measures of related phenomena. Maximum temperature produced a negative correlation with scrub typhus incidence and elevation produced a positive correlation. Positive correlations were also detected between scrub typhus incidence and NDVI, rainfall and the proportion in agriculture. Increasing population density was found to have a protective effect against scrub typhus, with very few cases being detected in urban settings. In terms of land cover, several land cover types were significantly correlated with the outcome: administrative areas with a higher proportion of crop-vegetation and forest had increased incidence of scrub typhus, whereas administrative areas with a high proportion of cropland had lower incidence. Increasing average income was also a protective factor for scrub typhus, which is likely to be related to population density, because densely populated, urban areas tend to have higher average incomes and also a smaller proportion in agriculture.

#### *Non-spatial Poisson regression analysis*

The final multivariate non-spatial regression model (the model with the lowest AIC) included the following covariates: proportion in agriculture; crop-vegetation; average income; and elevation (included as a quadratic expression). See Table 2 for the coefficients and p-values for each covariate in the final model. Adjusting for the effects of the other covariates, but not accounting for spatial autocorrelation in the data, the proportion in agriculture produced a negative correlation with scrub typhus incidence, which was opposite to the association seen during univariate analysis. Average income was also negatively correlated with the outcome, whilst increasing crop-vegetation was demonstrated to be a risk factor. Incidence of scrub typhus increased with increasing elevation, but at a

diminishing rate. A Moran's *I* test detected statistically significant spatial autocorrelation in the residuals from the multivariate, non-spatial Poisson regression model ( $P < 0.005$ ), indicating that the covariates included in the non-spatial model did not fully explain the spatial variation in scrub typhus. Thus, a spatial Poisson regression analysis was fitted, using the covariates from the best fitting non-spatial regression analysis, to address the spatial autocorrelation in the residuals.

#### *Spatial Poisson regression analysis*

The multivariate model containing fixed effects and both spatially correlated and uncorrelated random effects had the lowest DIC: 1,417.7 compared with 1,421.1 for the model with fixed effects and spatially correlated random effects; 1,473.8 for the model with fixed effects and uncorrelated random effects and 3,673.7 for the model with only fixed effects. See Table 3 for model coefficients from the full model with fixed effects and both spatially correlated and uncorrelated random effects. It was not possible to fit elevation as a quadratic term in the final regression analysis: this may have been caused by the quadratic term and the random effects competing for the same variation in the outcome measure. The model results demonstrated statistically significant positive correlations (95% CrI for the RR does not cross 1) with scrub typhus for both crop-vegetation and elevation. Average income was negatively associated with scrub typhus incidence; this was of marginal significance following inclusion of the random effects (95% CrI for the rate ratio [RR] included 1). The proportion in agriculture lost statistical significance in the final model. Fig. 3 illustrates the smoothed relative risks for scrub typhus within administrative areas of Taiwan, after having accounted for fixed and random effects. This shows a high risk of scrub typhus in the mountainous, rural areas in the eastern and central areas of the

Table 2. Final parameter estimates from multivariate model including fixed effects and spatially correlated and uncorrelated random effects.

| Variable                                       | Posterior mean (SD <sup>a</sup> ) | Rate ratio (95% CrI <sup>b</sup> ) |
|--|-----------------------------------|------------------------------------|
| Intercept                                      | -0.41 (0.05)                      |                                    |
| Average income (10 NT\$ per year increase)     | -0.006 (0.006)                    | 0.99 (0.98-1.00)                   |
| Crop-vegetation* (0.01% increase)              | 0.02 (0.008)                      | 1.02 (1.01-1.04)                   |
| Elevation* (100 m increase)                    | 0.20 (0.02)                       | 1.22 (1.18-1.26)                   |
| Proportion in agriculture (10% increase)       | 0.0004 (0.04)                     | 1.00 (0.92-1.08)                   |
| Uncorrelated random effect (precision)         | 0.63 (0.10)                       |                                    |
| Spatially correlated random effect (precision) | 16.30 (8.75)                      |                                    |

<sup>a</sup>Standard deviation; <sup>b</sup>credible interval; \*denotes statistical significance in the final Bayesian model (e.g. the 95% CrI for the rate ratio does not include 1). Note that the scale differs between variables.

island, low risk of disease in western areas, and some areas of medium risk to the far north and south-central areas of the island. Visual assessment of the spatial random effects (Fig. 4) demonstrates areas of high risk in the south east of Taiwan and areas of low risk on the west and north-west coasts that are not accounted for by the covariates in the model.

## Discussion

Quantifying the effect of climatic, environmental and socioeconomic variables on the incidence of disease provides enhanced understanding of disease epidemiology and spatial distributions, allowing the tailoring of disease control initiatives. The spatial analysis presented here indicates that scrub typhus incidence increases with both elevation and the proportion of land that is crop-vegetation mosaic, and a higher risk of scrub typhus occurs in the mountainous, rural areas in the eastern and central areas of Taiwan, while a lower risk of disease exists in the western part of the island.

Scrub typhus frequently occurs in habitats characteristic of transitional vegetation, such as forest edges, shrub-covered margins along streams or lands left untended after forest clearance/agricultural practices (Traub and Wisseman, 1974). Abandonment of agricultural cultivation creates ideal habitats for chiggers because crop remnants and natural vegetative regrowth provide the food and shelter necessary for proliferation of rodent hosts; secondary vegetation also helps provide moist terrain essential for chigger survival (Traub and Wisseman, 1974; Muul et al., 1977; Kuo et al., 2012). In this study, based on univariate Poisson regression, scrub typhus incidence in Taiwan was found to increase with the proportion of land covered with crop-vegetation mosaics and forest, while it decreased with the proportion of land covered

with cropland although it was not correlated with areas of forest-grass-shrub mosaics. Crop-vegetation was found to have the most significant correlation with scrub typhus incidence, which is likely due to the fact that crop-vegetation can provide favourable habitats for rodents by providing easy access to abundant food from nearby vegetative shelters (similarly to abandoned fields). Compared with natural habitats, more frequent human activities in cultivated fields and

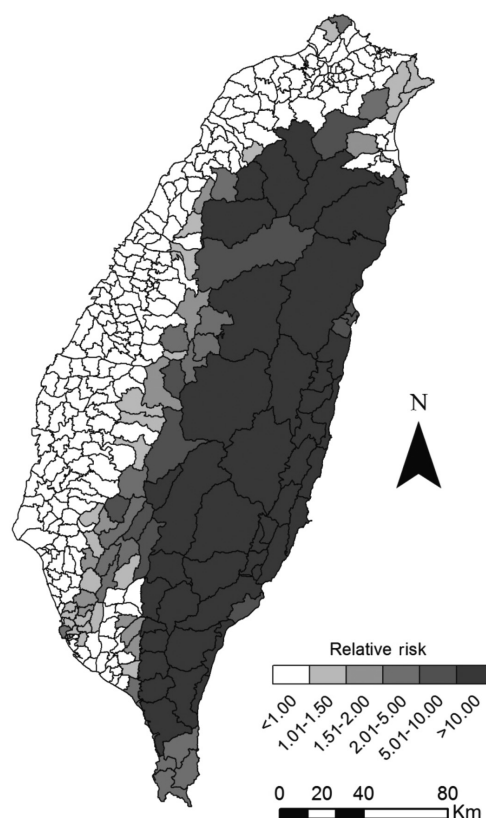


Fig. 3. Relative risk of scrub typhus within administrative areas of Taiwan, accounting for covariates and random effects.

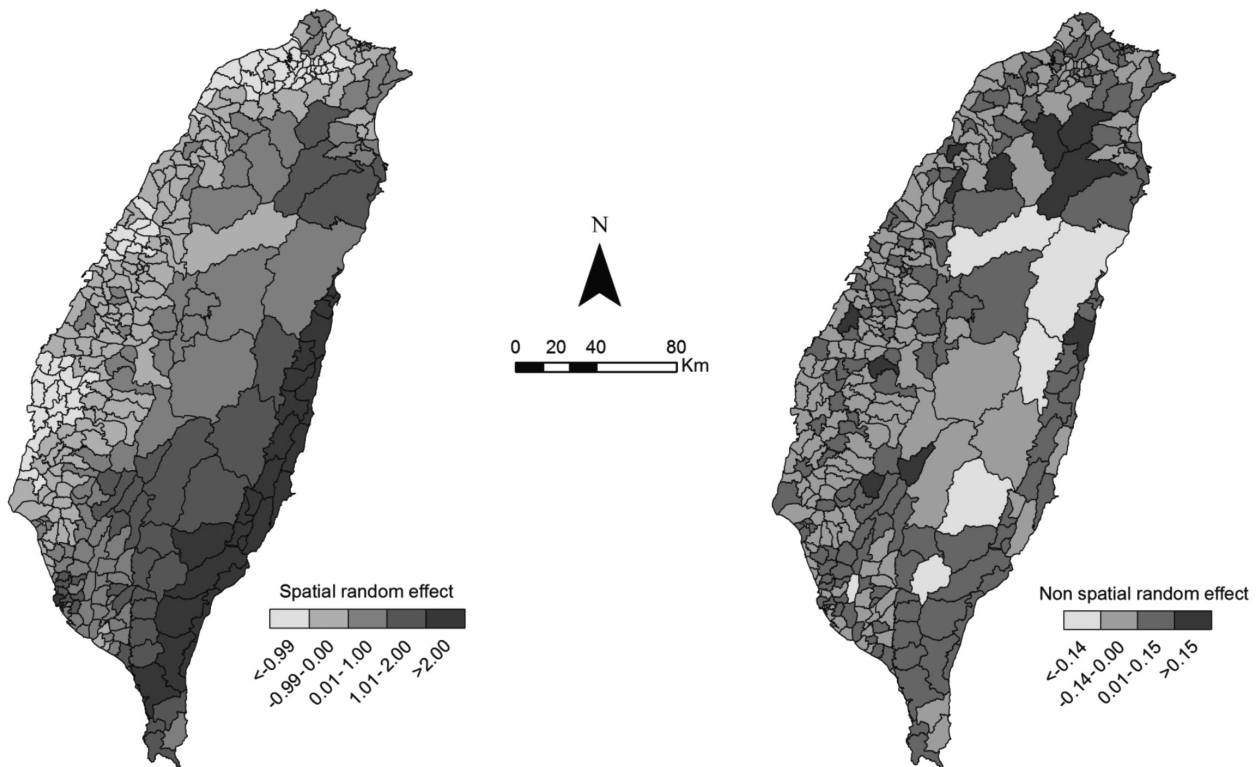


Fig. 4. Spatial (left map) and non-spatial (right map) random effects for scrub typhus.

crop-vegetation mosaics may also contribute to increased contact between humans and chiggers and, thus, result in increased transmission of scrub typhus to humans. This agrees with our previous findings that concurrence of higher NDVI (i.e. more vegetative growth) and higher proportion of farmers working in dry fields would lead to increased risk of scrub typhus (Kuo et al., 2011). In comparison, a negative correlation was seen between scrub typhus incidence and croplands in the univariate analysis, suggesting that contact between humans and chiggers is less frequent in areas that are more intensively farmed, possibly due to unsuitable habitat for trombiculid mites and their rodent hosts.

The significance of crop-vegetation for scrub typhus infection indicates that public health should be considered during land use policy decision making. In Taiwan, this land cover type tends to occur in less populated and mountainous areas (Fig. 2). Agricultural cultivation established in these frontiers is more likely to be abandoned later due to demographic and socioeconomic changes that render farming in these areas not feasible or profitable any more (Kuo et al., 2012). Once untended, these fields are readily reclaimed by surrounding vegetation, followed by colonization by rodents. The legacy of frontier cultivation in Taiwan

may, therefore, be increased risk of scrub typhus infection. The association of frontier colonization with outbreaks of scrub typhus has also been documented in several other countries (Traub and Wisseman, 1974), although our study is the first one to quantitatively investigate such an association, suggesting that, similar to other vector-borne diseases, land use policy that encourages agricultural practices in the frontiers should also contemplate the potential for elevated burdens of scrub typhus infection and other vector-borne diseases (Walsh et al., 1993; Singer and de Castro 2001).

The significance of elevation within the final model may be due to a combination of factors: land cover, climatic variables and socioeconomic factors. Several of the covariates studied vary across Taiwan in tandem with elevation (Lee et al., 2004): high elevation areas are less accessible and more difficult to cultivate, leading to lower population densities, higher NDVI values, lower socioeconomic status and differences in land cover and land use patterns. In addition, temperatures are lower and rainfall higher in high elevation areas. Thus, elevation may be acting as a proxy for a combination of these factors rather than as a significant correlate in terms of its own impact on disease transmission patterns. However, given the data available, it is



not possible to unravel the separate effects of these variables. Administrative areas with increased average incomes had lower risk of scrub typhus (although this was only of marginal statistical significance in the final model), reinforcing our understanding of scrub typhus risk relating to socioeconomic status, as well as climatic and environmental variables.

The correlation between scrub typhus incidence and the proportion in agriculture was positive during the univariate analysis, but this correlation was reversed in the non-spatial multivariate analysis and the covariate lost statistical significance in the final, spatial model. It may be expected that areas with a higher proportion of the population in agriculture would have a higher risk of scrub typhus, but the inclusion of land cover variables may already account for this effect (i.e. the effect of the proportion in agriculture on scrub typhus incidence changes when controlling for the effect of land cover). Moreover, the inclusion of crop-vegetation in the multivariate analysis suggests that the highest incidence occurs not in areas that are intensively farmed, but in areas where there is a transition from agricultural land cover to more natural habitats, which do not necessarily coincide with the areas with a high proportion of the population engaged in farming.

Univariate Poisson regression analysis indicated that all the variables assessed (excluding forest-grass-shrub) were significantly correlated with the incidence of scrub typhus in Taiwan during the study period. The highest incidence of disease occurred in rural areas, which tend to be at higher elevation and, thus, with lower average temperatures and higher annual rainfall. These areas also tend to be predominantly forest cover, which leads to high values of NDVI. The relatively large number of statistically significant covariates, along with collinearity between several of the variables, caused difficulties in the separation of variable effects and the selection of the best (or most plausible) set of covariates for the final model. Non-spatial Poisson regression analysis was used to select the best fitting model, but alternative model-selection strategies may have resulted in a different combination of covariates. However, examination of all the potential covariates and interpretation of their correlation with scrub typhus incidence helps to build a comprehensive picture of the spatial epidemiology of scrub typhus within Taiwan, which can be of assistance for the future development and targeting of control.

The presence of unaccounted for spatial variation in the data was illustrated by visualisation of the spatial random effects (Fig. 4). Areas with high values for the

spatial random effects in the south east of Taiwan indicated high incidence within these areas, which was not fully accounted for, and low value clusters on the western coast of Taiwan indicated areas of low incidence that were not fully explained by the model. The random effects allowed us to counteract the residual variation due to spatially heterogeneous variables not included in the final model; for example population density, which was found to be statistically significant in the univariate analysis, but was not included as a covariate in the final multivariate model.

The analysis presented made use of Bayesian inference. Although Bayesian methods can be computationally intensive and require significant start-up costs in terms of training and understanding, they offer a robust framework for the modelling of hierarchical and auto-correlated datasets and allow the incorporation of uncertainty in unknown model parameters (Best, 2005). Prior selection, which is required for Bayesian inference, can influence model outputs: here, non-informative priors were used to avoid bias. These results should be viewed with consideration of the research limitations. Due to the varying sizes of the administrative areas used in the analysis, the values used for several of the covariates (e.g. average elevation, average income) have different levels of accuracy (i.e. the elevation within a small administrative area will have a smaller variance than the elevation within a large administrative area): these differences may have an impact upon parameter estimation and significance levels. In addition, the spatial referencing of the scrub typhus case data was carried out based on the presumed location that infection occurred, although it can be difficult to ascertain the precise location that an infection was acquired. It is not possible to assess the reliability of these locations, but the aggregation of case counts to administrative areas will help to negate any imprecision in geo-referencing.

## Conclusions

The application of Poisson regression, including spatially and non-spatially correlated random effects, has highlighted the importance of a range of factors on the spatial distribution of scrub typhus within Taiwan. This research has provided evidence of increased scrub typhus risk within high elevation areas with land which contains a large proportion of crop-vegetation mosaics. In Taiwan, these transitional land cover areas tend to occur in less populated and mountainous areas, following the frontier establishment and subsequent partial abandonment of agricultural cultivation,

due to demographic and socioeconomic changes. Future land use policy decision-making should ensure that potential public health outcomes, such as modified risk of scrub typhus, are considered. Further analysis, incorporating a temporal aspect, is planned to provide further understanding of the environmental and climatic drivers of scrub typhus distributions. The investigation of seasonal patterns of infection may go some way to unravelling the separate effects of collinear climatic and environmental variables on scrub typhus incidence in Taiwan.

### Ethical statement

No individual case records or personally identifiable information were used as part of this study: disease records were aggregated to administrative areas. This study and the use of the data have been approved by the University of Southampton Research Ethics Committee.

### Acknowledgements

This study was supported by the Medical Research Council (NAW, PMA – project MR/J012343/1). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

### References

- Allan BE, Keesing F, Ostfeld RS, 2003. Effect of forest fragmentation on Lyme disease risk. *Conserv Biol* 17, 267-272.
- Arinaminpathy N, McLean AR, Godfray HCJ, 2009. Future UK land use policy and the risk of infectious disease in humans, livestock and wild animals. *Land Use Policy* 26, S124-S133.
- Best N, Richardson S, Thomson A, 2005. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res* 14, 35-59.
- European Space Agency, 2008. Globcover land cover map.
- Gentry JW, Phang OW, Manikumar C, 1977. The effects of rainfall on trombiculid (Acarina: trombiculidae) larval populations in peninsular Malaysia. *Southeast Asian J Trop Med Public Health* 8, 217-220.
- Jackson LE, Hilborn ED, Thomas JC, 2006. Towards landscape design guidelines for reducing Lyme disease risk. *Int J Epidemiol* 35, 315-322.
- Johnson BJ, Munafa K, Shappell L, Tsipoura N, Robson M, Ehrenfeld J, Sukhdeo MVK, 2012. The roles of mosquito and bird communities on the prevalence of West Nile virus in urban wetland and residential habitats. *Urban Ecosyst* 15, 513-531.
- Kawamura A, Tanaka H, Takamura A, 1995. *Tsutsugamushi disease: an overview*. Tokyo: University of Tokyo Press.
- Kéry M, 2010. *Introduction to WinBUGS for ecologists*. Burlington: Academic Press.
- Kuo CC, Huang JL, Ko CY, Lee PF, Wang HC, 2011. Spatial analysis of scrub typhus infection and its association with environmental and socioeconomic factors in Taiwan. *Acta Trop* 120, 52-58.
- Kuo CC, Huang JL, Shu PY, Lee PL, Kelt DA, Wang HC, 2012. Cascading effect of economic globalization on human risks of scrub typhus and tick-borne rickettsial diseases. *Ecol Appl* 22, 1803-1816.
- Lambin EF, Tran A, Vanwambeke SO, Linard C, Soti V, 2010. Pathogenic landscapes: Interactions between land, people, disease vectors, and their animal hosts. *Int J Health Geogr* 9, 54.
- Lawson AB, 2009. *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. Boca Raton: CRC Press.
- Lee PF, Ding TS, Hsu FH, Geng S, 2004. Breeding bird species richness in Taiwan: distribution on gradients of elevation, primary productivity and urbanization. *J Biogeogr* 31, 307-314.
- Lee PF, Liao CY, Lee YC, Pang YH, Fu WH, Chen SW, 1997. An ecological and environmental GIS database for Taiwan.
- Lee YS, Wang PH, Tseng SJ, Ko CF, Teng HJ, 2006. Epidemiology of scrub typhus in eastern Taiwan, 2000-2004. *Jpn J Infect Dis* 59, 235-238.
- Muul I, Lim BL, Walker JS, 1977. Scrub typhus infection in rats in four habitats in Peninsular Malaysia. *Trans R Soc Trop Med Hyg* 71, 493-497.
- Ogawa M, Hagiwara T, Kishimoto T, Shiga S, Yoshida Y, Furuya Y, Kaiho I, Ito T, Nemoto H, Yamamoto N et al., 2002. Scrub typhus in Japan: epidemiology and clinical features of cases reported in 1998. *Am J Trop Med Hyg* 67, 162-165.
- Olson JG, Bourgeois AL, 1979. Changing risk of scrub typhus in relation to socioeconomic-development in the Pescadores Islands of Taiwan. *Am J Epidemiol* 109, 236-243.
- Patz JA, Daszak P, Tabor GM, Aguirre AA, Pearl M, Epstein J, Wolfe ND, Kilpatrick AM, Foutopoulos J, Molyneux D et al., 2004. Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence. *Environ Health Perspect* 112, 1092-1098.
- Rosenberg R, 1997. Drug-resistant scrub typhus: paradigm and paradox. *Parasitol Today* 13, 131-132.
- Seong SY, Choi MS, Kim IS, 2001. *Orientia tsutsugamushi* infection: overview and immune responses. *Microbes Infect* 3, 11-21.
- Singer BH, de Castro MC, 2001. Agricultural colonization and malaria on the Amazon frontier. *Ann NY Acad Sci* 954, 184-222.
- Traub R, Wisseman CLJ, 1974. The ecology of chigger-borne rickettsiosis (scrub typhus). *J Med Entomol* 11, 237-303.
- Tsai PJ, Yeh HC, 2013. Scrub typhus islands in the Taiwan area and the association between scrub typhus disease and forest land use and farmer population density: geographically weighted regression. *BMC Infect Dis* 13, 191.

- Vallee J, Thaojaikong T, Moore CE, Phetsouvanh R, Richards AL, Souris M, Fournet F, Salem G, Gonzalez JPJ, Newton PN, 2010. Contrasting spatial distribution and risk factors for past infection with scrub typhus and murine typhus in Vientiane City, Lao PDR. *PLoS Negl Trop Dis* 4, e909.
- Vittor AY, Pan W, Gilman RH, Tielsch J, Glass G, Shields T, Sanchez-Lozano W, Pinedo VV, Salas-Cobos E, Flores S et al., 2009. Linking deforestation to malaria in the Amazon: characterization of the breeding habitat of the principal malaria vector, *Anopheles darlingi*. *Am J Trop Med Hyg* 81, 5-12.
- Walsh JF, Molyneux DH, Birley MH, 1993. Deforestation - effects on vector-borne disease. *Parasitology* 106, S55-S75.
- Zhao SQ, Peng CH, Jiang H, Tian DL, Lei XD, Zhou XL, 2006. Land use change in Asia and the ecological consequences. *Ecol Res* 21, 890-896.