Geospatial Health 2021; volume 16:879

Spatial patterns of *Plasmodium vivax* transmission explored by multivariate auto-regressive state-space modelling - A case study in Baoshan Prefecture in southern China

Jinxin Zheng,[1-4] **Benyun Shi,**[5] **Shang Xia,**[1-4] **Guojing Yang,**[6,7,8] **Xiao-Nong Zhou**[1-4]

[1]National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai, China; [2]Key Laboratory of Parasite and Vector Biology, National Health Commission, Shanghai, China; [3]National Center for International Research on Tropical Diseases, Ministry of Science and Technology, Shanghai, China; [4]Chinese Center for Tropical Diseases Research, Shanghai, China; [5]School of Computer Science and Technology, Nanjing Tech University, Nanjing, Jiangsu, China; [6]Hainan Medical University, Laboratory of Tropical Environment and Health, Haikou, Hainan, China; [7]Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute; [8]University of Basel, Basel, Switzerland

Abstract

The transition from the control phase to elimination of malaria in China through the national malaria elimination programme has focussed attention on the need for improvement of the surveillance-response systems. It is now understood that routine passive surveillance is inadequate in the parasite elimination phase that requires supplementation by active surveillance in foci where cluster cases have occurred. This study aims to explore the spatial clusters and temporal trends of malaria cases by the multivariate auto-regressive state-space model (MARSS) along the border to Myanmar in southern China. Data for indigenous cases spanning the period from 2007 to 2010 were extracted from the China's Infectious Diseases Information Reporting Management System (IDIRMS). The best MARSS model indicated that malaria transmission in the study area during 36 months could be grouped into three clusters. The estimation of malaria transmission patterns showed a downward trend across all clusters. The proposed methodology used in this study offers a simple and rapid, yet effective way to categorize patterns of foci which provide assistance for active monitoring of malaria in the elimination phase.

Correspondence: Xiao-Nong Zhou, National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai 200025, China. E-mail: zhouxn1@chinacdc.cn

Key words: Malaria; *Plasmodium vivax*; Indigenous case; State space model, Foci; cluster; China.

Acknowledgements: we would like to thank the support of Disease Prevention and Control Center of Yunnan Province, China, and Chinese Center for Disease Control and Prevention in Shanghai and all the staff members in data collection and investigation. We are also grateful to the study participants for their commitment.

Funding: this research was funded by National Natural Science Foundation of China (*Grant No. 81402760, 81573261, 81502858*), as well as a grant from Natural Science Foundation of Jiangsu Province, China (*Grant No. BK20161563*) also with the Open Project of Key Laboratory of Parasite and Vector Biology, China Ministry of Health (*WSBKFKT-201804*). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest: the authors declare no potential conflict interests.

Received for publication: 9 March 2020.
Revision received: 21 August 2020.
Accepted for publication: 21 August 2020.

Geospatial Health 2021; 16:879
doi:10.4081/gh.2021.879

Introduction

Malaria is a life-threatening disease caused by *Plasmodium* spp. parasites that are transmitted through the bite of infected female *Anopheles* spp. mosquitoes. In 2019, the World Health Organization (WHO) estimated the worldwide disease burden at nearly 228 million cases and 405,000 deaths (WHO, 2019). In order to control and eliminate malaria globally, various projects or programmes have been proposed, such as the Global Malaria Programme (Whittaker *et al.,* 2014), the Asia Malaria Elimination Network (Gosling *et al.,* 2012), and the Strategy for Malaria Elimination in the Greater Mekong Sub-region (Cui *et al.,* 2012) and others. Ten years ago, the National Malaria Elimination Action Plan (NMEAP) was launched by the Ministry of Health (MoH) in the People's Republic of China (MoH, 2010). Thanks to effective strategies implemented by the Chinese Government, the malaria incidence rate has decreased significantly since then. NMEAP included disease management and vector control, which have been confirmed to be powerful and co-effectiveness measures (Yin *et al.,* 2014; Zhou *et al.,* 2015b; Hu *et al.,* 2016).

Most malaria cases are reported from south-western China, especially along the China-Myanmar border in Yunnan Province (Wang *et al.,* 2014; Shi *et al.,* 2015; Xu *et al.,* 2016). Indeed, from 2013 to 2017, up to 90% of all indigenous malaria cases in China originated in these areas (Zhou *et al.,* 2011; Li *et al.,* 2016; Feng *et al.,* 2018; Zhang *et al.,* 2018). For examples, 3078 malaria cases were reported across the country in 2014, 2% of which were indigenous and had occurred in the China-Myanmar border areas (Li *et al.,* 2016). Thus, malaria transmission here remains a big challenge for NMEAP, which aims to free the country of malaria by 2020.

Active surveillance is an effective approach in the malaria elimination phase (Zhang *et al.,* 2017; Zhou *et al.,* 2015b). However, implementation of this kind of surveillance in the endemic areas is a crucial matter, especially in the areas characterized by complicated environmental settings. The geographic variations of malaria transmission in China-Myanmar border areas have shown an underlying spatial pattern of malaria incidence providing a reference to design local surveillance strategies (Shi *et al.,* 2015). Along with this point, it is necessary to explore the spatial patterns of malaria transmission so as to provide evidence-based data that can improve the sensitivity of the surveillance-response system in the elimination phase. Since 2004, the Infectious Diseases Information Reporting Management System (IDIRMS), established by the Chinese Center for Disease Control and Prevention (China CDC), different infectious diseases have been closely examined and scrutinized in the county health care institutions and then rapidly reported to IDIRMS. This well-run data collection provides a good opportunity to analyze the dynamic patterns of malaria transmission, *e.g.*, by mathematic modelling, which has provided a way to quickly estimate or predict transmission risks of infectious diseases.

Ever since the availability of large datasets including times series on disease transmission, the auto-regressive integrated moving average (ARIMA) approach has been widely used to forecast the transmission pattens of vector-borne diseases (Wangdi *et al.,* 2010; Anwar *et al.*, 2016). These studies mainly focused on single time series (Wangdi *et al.,* 2010) rather than multiple ones from different sites. However, studies in this field have started to apply techniques, so far mainly used in artificial intelligence and game theory, for analytical investigations of multivariate time-series in fields, such as parasitology, ecology, economics and genetics (Valle and Clark, 2013; Bartocci and Lió, 2016; Wang, He *et al.,* 2016; Tolimieri *et al.,* 2017). The state-space is central in this kind of modelling. It represents all possible configurations of a system, where the coordinates are state variables that together describes the state of the system completely, which means that each point in the state-space corresponds to a different state of the system. As this kind of modelling uses abundant time series to balance limited data support, this strategy is suitable for dealing with small areas and populations. The multivariate auto-regressive state-space (MARSS) technique, originally proposed by Hinrichsen (2009) and further developed by Ward *et al.* (2010) and Holmes *et al.* (2012), provide a useful mathematical approach to work on multiple time series. Since this approach has also been applied to predict malaria transmission (Valle and Clark, 2013), we felt that it would be possible to attempt using MARSS for tracing malaria transmission in southern China. Baoshan, Yunnan Province, is located in south-western China near the China-Myanmar border. Due to the nature of the geological environment and mosquito species as well as frequent mobility of migrated population along the border, malaria is still a serious problem in this region. It has consistently shown the highest number of malaria cases due to *Plasmodium vivax* in the whole country during 2007-2013 (Xu *et al.,* 2016). In this study, we attempted to develop a model providing a reference for decision-making with regard to surveillance policy aimed at eliminating malaria in China-Myanmar border areas by addressing two questions: i) is active surveillance of malaria in the Baoshan region consistent with the strategies applied in the geographical administrative regions and the disease foci?; and ii) do the Baoshan counties represent different spatial transmission patterns than seen elsewhere?

Materials and methods

Study area

Malaria transmission was studied in Baoshan, one of the prefectures in Yunnan Province bordering Myanmar, which has an area of 19,637 km^2 and a population of 2.5 million based on the 2010 census. It has a tropical climate with abundant precipitation, averaging 700-1200 mm per year and a seasonal temperature fluctuation with peaks in May-August. There are five administrative counties under Baoshan, namely Changning, Longling, Longyang, Shidian and Tengchong (Figure 1). Malaria cases appear all year around in Baoshan Prefecture accounting for the majority of malaria cases in Yunnan Province and China as a whole before 2010, especially in Tengchong County near the China-Myanmar border with high numbers of annual malaria cases, the highest in China.

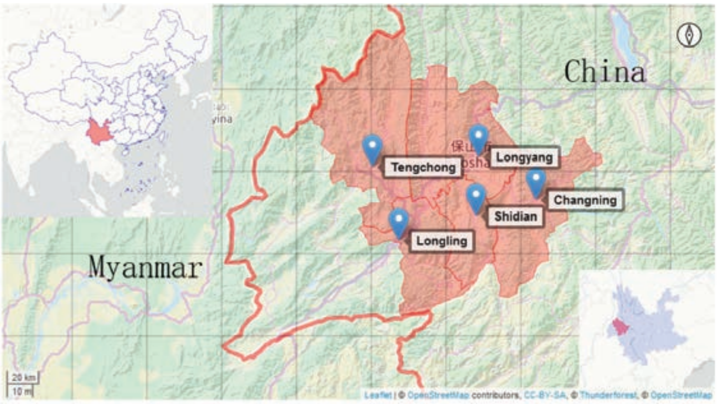

Figure 1. Map of malaria transmission in five counties in Baoshan Prefecture, Yunnan Province, China near Myanmar border.

Data collection

Information on indigenous *P. vivax* cases in the study area were provided by IDIRMS of China CDC. We aggregated indigenous *P. vivax* incidence data from January 2007 to December 2009 by patient residential addresses, which were distributed over all five counties of Baoshan Prefecture. To fit the MARSS model, we used time series data transformed into the monthly distribution of *P. vivax* cases.

State-space model

The utilized the MARSS model as specified by Holmes *et al.* (2012), by which the total variance in the observation of time series is defined by partition into observation variance and process variance. In our study, the latter is the temporal variance of the malaria declining rate, while former includes sampling error and measurement error from data inaccurcies. The MARSS framework is given by two equations:

$$X_t = B_t X_{t-1} + u_t + w_t\ ,\ w_t \sim N[0,\ Q_t] \qquad (1)$$

$$Y_t = Z_t X_t\ + a_t + v_t\ , \qquad v_t \sim N[0,\ R_t] \qquad (2)$$

We used n to denote the number of discrete survey counties, and m to represent the number of unknown hidden states (foci clusters). For the hidden states in Eq (1), X_t is a matrix that denotes the vector of length m representing the true observation indigenous cases in year t; B_t a $m\times m$ matrix that allows the density dependence; u_t a $m\times 1$ vector of hidden states of the declining rates; and w_t a vector of length m representing the process errors in year t. We assumed that the process errors can be correlated and specified from a multivariate normal distribution with a zero mean of a $m\times m$ matrix Q_t. In Eq (2), Y_t is the n element vector of observed malaria cases at time t, Z_t a $n\times m$ matrix that defines n observations related to the m hidden states. In general, the n sites could be any samples from the m hidden sates. The $n\times 1$ vector specifies the bias between those sites. The observation errors at time t are denoted by v_t, the same approach to address process error in Eq (1), which is an n vector of serially uncorrelated disturbances with zero mean of an $n\times n$ matrix R_t. When using our data of all malaria cases from the five counties (the observed time series), we wanted to fit a model with one hidden state. All observed time series were sampled from this unique focus cluster. There was one process error variance with different observation variance. The observed malaria cases were independent of each other. Mathematically the fitted MARSS model is described by the following matrices:

$$x_t = x_{t-1} + u + w_t, w_t \sim N(0,q)$$

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \\ v_{4,t} \\ v_{5,t} \end{bmatrix}, v_t \sim MVN(0, \begin{bmatrix} r_1 & 0 & 0 & 0 & 0 \\ 0 & r_2 & 0 & 0 & 0 \\ 0 & 0 & r_3 & 0 & 0 \\ 0 & 0 & 0 & r_4 & 0 \\ 0 & 0 & 0 & 0 & r_5 \end{bmatrix})$$

In this formula, the v_t comes from a multivariate normal distribution (MVN) with variance-covariance matrix.

Model estimation and selection

Our goal was to find all possible foci clusters in the dataset, then evaluate the best fitted model. Assuming that the observation errors are independent within all the sites, we fitted a series of models for all of sites in Baoshan, for each model the process error and reduction rates were set differently within each cluster. Our model selection was based on the value of Akaike's information criterion (AIC) as given by (Koc and Bozdogan, 2015). It offers an estimation of the relative information lost when a given model is used, and when the dataset is small, only a small sample modification can be supported (Holmes *et al*., 2012). We used the corrected AIC (AICc) that is adjusted to small datasets. The maximum likelihood parameter is used to provide fully parametric bootstraps, when the outputs of parameters are sampled from a numerically estimated Hessian matrix (Holmes *et al.,* 2012). In our study, we evaluated the model with the Shapiro-Wilk test (Royston, 1982) to identify the standardized residuals, checking the autocorrelations for the structure left in the residuals, which computes an estimate of the auto-correlation function (ACF) of the multivariate time series (Aryee *et al.,* 2018). Due to the complexity of the multivariate models and the small sample sizes of the malaria cases, we used a 90% confidence interval to evaluate the MARSS simulated values in each cluster for estimation of the total trends and true patterns of malaria. Bootstrapping can be used to construct frequency condence intervals of the parameter estimates, and the MARSS model includes a semi-parametric bootstrap algorithm (Holmes *et al.,* 2012) that affects the model residuals, *i.e.* the innovations. This approach uses the maximum-likelihood MARSS parameters to simulate data from which bootstrap parameter estimates are obtained. All test and data were analysed in R 3.3.2 software with MARSS packages (version 3.9) as used by (Holmes *et al.,* 2012).

Results

The total number of the indigenous *P. vivax* cases in the study period from 2007 to 2009 showed nearly five times more indigenous cases in Tengchong than in any other county. The lowest case number ocurred in Changning County, where there was 10 cases as compared to none at all in 2009. The number of malaria cases in the other three counties ranged from 0 to 32 within the three study years (Figure 2). A decreasing trend of indigenous *P. vivax* cases was observed in the Baoshan region in the period 2007-2009. Seasonal components varied among the five counties, such as an obvious periodicity pattern every year in Tengchong County, a peak from April to July occurred in Longning, Longyang and Shidian counties, while a climbing pattern with a peak from June to August was observed each year in Changning County. From the general survey of our research area, the map showed that more cases appeared in Tengchong, which is situated near the China-Myanmar border, while fewer cases happened in Changning situated further inside the Baoshan Prefecture and thus far from the border (Figure 1). Many foci cluster combinations were possible in our MARSS modelling. In this study, however, we only listed the top 20 best-fitted models according to the AICc values. The best-fitted model, which had the lowest AICc value (286.417), contained three hidden states (three foci clusters) (Table 1). In the models with the lowest AICc values, the malaria cases were all grouped into three clusters based on the hypotheses that all counties were sampled from different populations (Table 1). Tengchong represented cluster 1, Shidian was part of cluster 2, while the other three counties shared cluster 3 (Table 2). According to Table 2, each group had an unconstrained process error indicating that a group of counties within the same cluster had the same spatial pattern and that their dynamics were homogeneous. Meanwhile, the observation errors were found to be unequal with a diagonal matrix. Tengchong's observation error was lower (0.0135), while Changning had a largeer error (0.3125) than all the other counties. As observed, the number of indigenous cases decreased in all of five counties, with negative reduction rates in all three clusters. The reduction rate in cluster 3 (–0.0537) decreased faster than that in the other two clusters suggesting the malaria infection is coming down in those counties. We evaluated the MARSS model by examination of the residual values, and comparison of the malaria indigenous cases between simulated and observed numbers.The simulation for the three clusters are shown in Figure 3. Tengchong was best-fitted to cluster 1 and cluster 2 did not extend beyond the

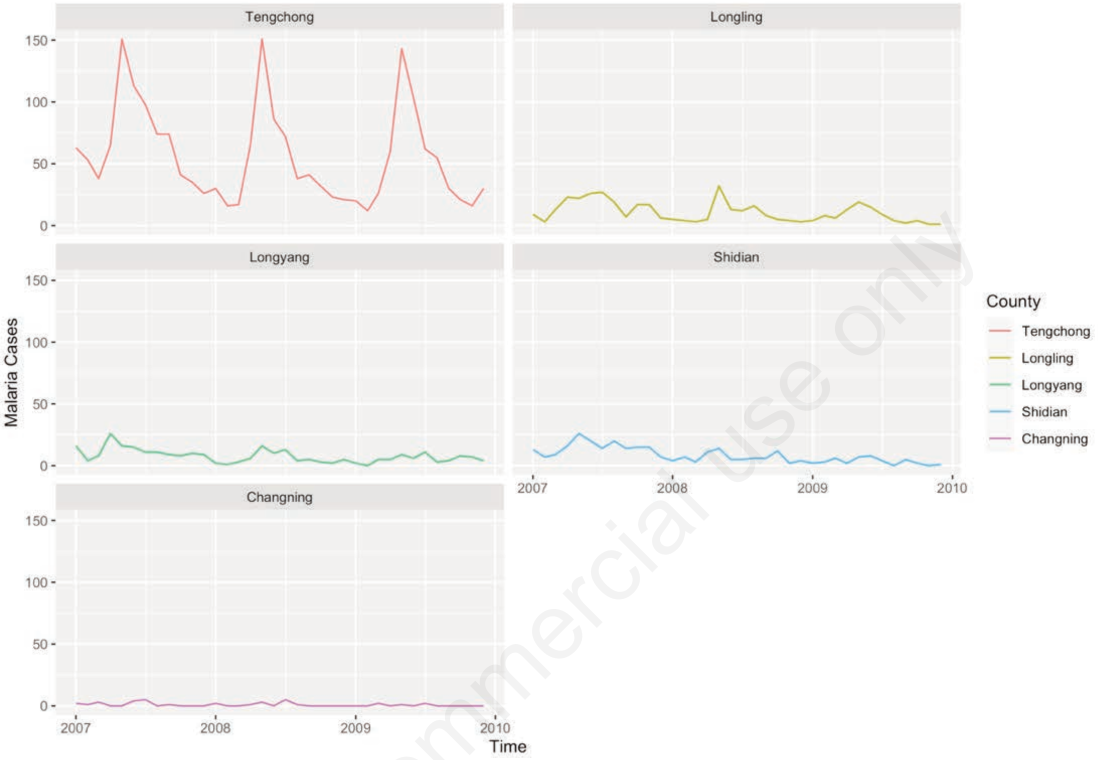

Figure 2. Monthly reported local *Plasmodium* vivax cases from five sites of Baoshan area, Yunnan Province, China.

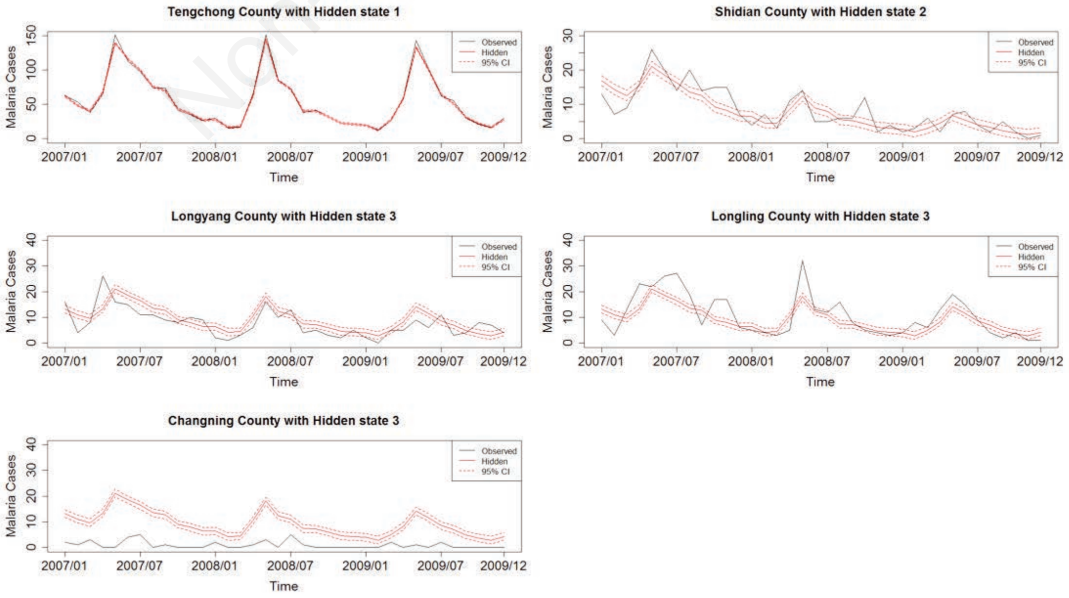

Figure 3. Estimation of best-fitted model for three clusters among the five sites and their 90% confidence intervals.

borders of Shidian, while cluster 3 covered parts of all remaining three sites (Longling, Changning and Longyang). The fitted simulation in the three foci clusters was different from each other, with cluster 3 over-estimated in Changning which had the lowest number of indigenous cases. The 90% confidence interval in each foci cluster contained more than 90% of the malaria observation in the model simulations.

Discussion

This study attempted to apply a time series model to detect the clusters of foci by using malaria surveillance data in south-western China near the Myanmar border in order to improve the surveillance sensitivity by targeted monitoring. China launched the malaria elimination programme in 2010, and no indigenous cases were reproted in 2017 in the whole country based on passive surveillance (Feng *et al.,* 2018), which is the main strategy to monitor malaria in China. However, there is a need to locate high-risk sites precisely and focus on the local level. For example, the re-establishment of high-risk malaria transmission in south-western China (Garros *et al.,* 2006; Zhang *et al.,* 2017; Yu *et al*., 2013; Li *et al.,* 2015) remains to this day due to specific, climactic factors and proximity to endemic countries in South. Therefore, it is essential for local programme managers to make decisions and develop policies to further improve the sensitivity of the surveillance-

Table 1. The performance of 20 models ranked by Akaike's information criterion corrected based on different hypotheses with respect to the malaria cases.

Model	Q[a]	R[b]	U[c]	LogLik[d]	AICc[e]	AIC[f]	Parameters[g]	M[h]	Hypothesis[i]
1	1	1	1	–121.8338	286.4176	281.6676	19	3	T-S-LYC
2	1	1	1	–122.4028	287.5555	282.8055	19	3	S-C-TLY
3	1	1	1	–122.7177	288.1853	283.4353	19	3	T-LS-YC
4	1	1	1	–122.8298	288.4095	283.6595	19	3	S-TL-YC
5	1	1	1	–118.3369	292.4157	284.6738	24	4	T-L-S-YC
6	1	1	1	–127.4923	287.9114	284.9846	15	2	SC-TLY
7	1	1	1	–127.5118	287.9504	285.0236	15	2	T-LYSC
8	1	1	1	–127.5694	288.0657	285.1389	15	2	S-TLYC
9	1	1	1	–123.9823	290.7146	285.9646	19	3	T-LY-SC
10	1	1	1	–118.9946	293.7311	285.9891	24	4	L-S-C-TY
11	1	1	1	–124.0113	290.7725	286.0225	19	3	L-S-TYC
12	1	1	1	–124.1037	290.9574	286.2074	19	3	C-TY-LS
13	1	1	1	–124.2693	291.2885	286.5385	19	3	T-C-LYS
14	1	2	1	–128.5523	290.0315	287.1047	15	3	S-C-TLY
15	1	1	1	–128.6523	290.2313	287.3045	15	2	TY-LSC
16	1	1	1	–124.8893	292.5287	287.7787	19	3	S-TY-LC
17	2	1	1	–129.8895	290.3244	287.779	14	2	LS-TYC
18	1	1	1	–128.9868	290.9005	287.9737	15	2	TL-YSC
19	1	1	1	–125.0074	292.7648	288.0148	19	3	L-TY-SC
20	2	1	1	–129.2122	291.3511	288.4243	15	3	S-TC-LY

[a]Process errors with 1 for independent variances and 2 for the same variance; [b]observation errors with 1 for independent variances and 2 for same variance; [c]reduction rate with 1 for unequality among the data; [d]Log-likelihood of the data conditioned on the model parameters; [e]Corrected Akaike information criterion; [f]Akaike information criterion; [g]Number of parameters in the model; [h]Hypothesis of hidden states among the five sites; [i]Counties grouped by hidden states. T, Tengchong; L, Longling; Y, Longyang; S, Shidian; C, Changning.

Table 2. Estimation of cluster growth rates, observation errors and process error in the five sites for the best-fitted model.

Cluster	County	Reduction rate	Observation error	Process error
1	Tengchong	–0.0223 (–0.1783; 0.4858)	0.0135 (0.0048; 0.0910)	0.2272 (0.0771; 0.3261)
2	Shidian	–0.0537 (–0.0909; 0.0165)	0.1517 (0.0884; 0.2320)	0.0509 (0.0003; 0.1010)
3	Longling	–0.0283 (–0.0240; 0.0753)	0.2229 (0.0570,0.1298)	0.0880 (0.0218; 0.1455)
	Longyang		0.2074 (0.1248; 0.3108)	
	Changning		0.3125 (0.1771; 0.4858)	

Bootstrapped 90% confidence interval for the reduction rate and the multivariate auto-regressive state-space model observation and process errors.

response systems by targeting resources to identify foci of risk through sentinel monitoring (Hay *et al.,* 2008).

Lately, no malaria cases have been reported from Baoshan Prefecture (Li *et al.*, 2015), which was confirmed in the 2017 surveillance report. However, transmission varies from place to place and our results showed that the best-fitted model was the three-cluster pattern in Baoshan region, which implies that the five counties near the Myanmar border carry a high risk for continued transmission. It is obvious from Figure 1 that Tengchong near the border had a higher risk of malaria occurrence, something that has already been pointed out in previous studies (Dong *et al.,* 2000). From the biological nature of malaria transmission, the geographical environment and climate in these areas represent intricate variables. In addition, vectors, such as *Anopheles dirus* and *An. minimus* do not respect political borders, which has recently been shown many times (Wang *et al.,* 2015b; Chen *et al.* 2017; Zhang *et al.*, 2018). Large numbers of *An. minimus, An.dirus* as well as *An.sinensis* populations have all been found in in Tengchong County (Dong, 2000; Wang *et al.,* 2015b). Their habitats are associated with rice fields and various breeding sites. If those vectors find suitable ecological conditions during the transmission season, there is a strong possibility of re-introduction of malaria through imported *P. vivax* cases (Zhou *et al.,* 2015a). Many reports show the cross-border activities play a major role in malaria transmission in the study region (Stoddard *et al.,* 2009; Xu *et al.,* 2016; Zhang *et al.,* 2016; Shi *et al.,* 2017), activities that also have varied from place to place and time to time, such as i) trade becoming more frequent in Tengchong County (Wang *et al.,* 2015a; Li *et al.*, 2016; Xu *et al.,* 2016); ii) increasing numbers of imported cases enter this area continually (Yin *et al.*, 2014); and iii) movements of people, *e.g*., labourer and farmers travelling across the border in particular months, leading to accumulated exposure opportunities (Yin *et al.*, 2014). Reports show that Tengchong and Longling generally harbour the highest number of malaria cases, which are ascribed shared borders with Myanmar (Wang *et al.*, 2016c) . Those living in other parts of Baoshan form relatively closed economic circumstances that result in fewer cross-border travel (Li and Yang, 2015).

We find that our work based on MARSS modelling of time series matrix data provides real-time transmission information including the detection of risk spatio-temporal clusters. However, our cluster measures are different from China's "1-3-7" surveillance-response, which report malaria cases within 1 day, conducts a case investigation within 3 days and in-depth investigation within 7 days (Zhou *et al.,* 2015b). This approach is based on cases reported in persons and sites, which means that if a case is identified, the "1-3-7" response is implemented limiting the number of cases in surrounding areas. Applying the MARSS model, however, spatiotemporal malaria clusters are instead searched at the regional level so that actions and surveillance can be focused at cluster regions when found. The whole region is thus considered and interventions can then be better directed, and the "1-3-7" surveillance-response more quickly targeted based on prepared resources and findings.

A number of studies have attempted to use biological data to define the structure of malaria transmission in the China-Myanmar border area. Firstly, *Anopheles* can easily transfer malaria from Myanmar to China or *vise versa* (Xu and Liu, 2016). The intensity depends mainly on density of the mosquitoes population and the local malaria prevalence, which is considered important transmission indicators in the border areas (Zhang *et al.,* 2016). Work on genetic diversity of the mosquito population in Yunnan Province (Feng *et al.,* 2017) has indicated that the vector gene flow is limited by geographical distance and geographical barriers (Liu *et al.,* 2016). Secondly, the mosquito insecticide resistance varies from area to area, which contributes to diversity with repect to malaria incidence; however this resistance does not always overlap in sites at each sides of the border, something that may be explained by gene diversities of malaria populations that has so far been studied in larger areas, such as Yunnan Province as a whole (Wang *et al.,* 2016b). Enlarged msquito genetics should be useful as the cross-border problems need to be solved. Further investigations of the different patterns of malaria transmission at the county level are thus warranted.

The incidence rate of malaria declined in Yunnan, but the rates fell differently from place to place. Such discrepancies across clusters in the study area are not surprising given the siuation near the Myanmar border compared to areas further away from the border (Figure 1). Although the overall annual number of indigenous cases have decreased gradually over the last few years, a pattern probably due to effective shared interventions implemented jointly in both China and Mynamar (Yin *et al.,* 2014; Zhou *et al.,* 2015b; Xu *et al.*, 2017; Feng *et al.,* 2018), we noticed differences of the reduction rate among the three clusters. However, even if the indigenous *P. vivax* cases in Tengchong tended down, there were still numerous cases in this county (Dong *et al.*, 2000; Shi *et al.,* 2015), which explains the good fitting results of the cluster simulations. Different locations had a different impact on malaria control, even if the total number of imported cases did not differ (Shi *et al.,* 2014), which could be the cause of the difference in the reduction rate in these sites. Overall, the three clusters showed a high heterogeneity but shared a decreasing trend. To be successful, active surveillance should contain foci related to all the spatial patterns discovered in the Baoshan area.

The MARSS model could analyze the inside structures within the multiple time series based on various spatio-temporal clusters. Our modeling approach relied on the hypothesis that the reported malaria cases were from one population structure or more than one population. However, the situation gets progressively more complicated as the number of time series increase. The number of permutations and combinations of many clusters from multiple sites rapidly becomes too big for listing in a simple table. This should be considered when using MARSS model. To dispose of the problem of the permutations produced by the algorithm used, Ward *et al.* (2010) subdivided 13 sites into several sub-populations based on previous knowledge on the biology of different species. Although this approach can provide an explanation for the hypotheses, it can also miss some information. Furthermore, the integrated nested laplace approximation (INLA) method (Costa *et al.*, 2011; Bivand *et al.,* 2015), based on the spatial temporal Bayesian model, is another way of applying effective computation to predict the spatial dynamics of infection diseases. In our study, the malaria data were collected from many sites and identification of the clusters by time series data was our major purpose. So there is a need to update this state space model about permutation and combination algorithm to get a real understanding of the malaria transmission patterns. On the other hand, the surveillance-response systems used were based on the population level, which offers baseline access data. Application of MARSS modelling on small-area surveillance data would produce more accurate active surveillance.

The limitations of the study are reflected in the following four aspects. The first challenge was that the data collection represented the county level and was based at only five restricted sites, which increased the risk of ending up with a comparatively large proportion of less precise data (Zhang *et al.,* 2018). Second, from the simulation part, a low number of positive malaria data could result in a poor simulations leading to the over-estimated results from some counties. Third, the spatial distributions of infections emphasized that the risk area does not coincide with the administrative regions. Fourth, risk factors based on demography, environment and socioeconomy were not considered in the modelling.

The purpose of this study was to estimate malaria cluster structures about Baoshan area by MARSS, which produces spatial patterns from time series data. This could be a direct way to obtaining a general overview of malaria foci clusters supported by strong surveillance-response systems, as China is now approaching the stage of malaria elimination. Therefore, although the aforementioned four limitations exist, the MARSS model is still the best way to identifiy the clusters in this area, which in turn leads the delivery of health resources more efficiently (Feng *et al.,* 2014).

Conclusions

Identification of focal clusters is important for the malaria elimination phase and critical for estimating the precise need for the health resources to be used for active surveillance. Applying the MARSS state-space model to several time series data at the county level made it possible to successfully estimate the reduction rate and the spatial patterns of indigenous malaria cases. Prevention of re-introduction of transmission into areas without any reported indigenous cases, requires processing data based on long-term monitoring by new methods. The approach investigated has provided a new, simple and direct tool to identify disease distributions capable of also predicting hidden states within trend parameters wich should help supporting long-term policies.

References

Anwar MY, Lewnard JA, Parikh S, Pitzer VE, 2016. Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. Malar J. 15:566.

Aryee G, Kwarteng E, Essuman R, Nkansa Agyei A, Kudzawu S, Djagbletey R, Owusu Darkwa E, Forson A, 2018. Estimating the incidence of tuberculosis cases reported at a tertiary hospital in Ghana: a time series model approach. BMC Public Health 18:1292.

Bartocci E, Lió P, 2016. Computational modeling, formal analysis, and tools for systems biology. PLoS Comput Biol 12:e1004591.

Bivand R, Gómez-Rubio V, Rue H, 2015. Spatial data analysis with R-INLA with some extensions. J Stat Softw 63:1-31.

Costa AC, Codeço CT, Honório NA, Pereira GR, N Pinheiro CF, Nobre AA, 2015. Surveillance of dengue vectors using spatiotemporal Bayesian modeling. BMC Med Inform Decis Mak. 15:93.

Chen T, Zhang S, Zhou SS, Wang X, Luo C, Zeng X, Guo X, Lin Z, Tu H, Sun X, Zhou H, 2017. Receptivity to malaria in the China-Myanmar border in Yingjiang County, Yunnan Province, China. Malar J 16:478.

Cui L, Yan G, Sattabongkot J, Chen B, Cao Y, Fan Q, Parker D, Sirichaisinthop J, Su XZ, Yang H, Yang Z, Wang B, Zhou G, 2012. Challenges and prospects for malaria elimination in the Greater Mekong Subregion. Acta Trop 121:240-5.

Dong X, 2000. [The malaria vectors and their ecology in Yunnan Province.] Chin J Parasit Dis Contr. [Article in Chinese]. Available from: http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZISC200002032.htm

Feng X, Huang L, Lin L, Yang M, Ma Y, 2017. Genetic diversity and population structure of the primary malaria vector Anopheles sinensis (Diptera: Culicidae) in China inferred by cox1 gene. Parasit Vectors 10:75. Erratum in: Parasit Vectors 10:120.

Feng XY, Xia ZG, Vong S, Yang WZ, Zhou SS, 2014. Surveillance and response to drive the national malaria elimination program. Adv Parasitol 86:81-108.

Feng J, Zhang L, Huang F, Yin JH, Tu H, Xia ZG, Zhou SS, Xiao N, Zhou XN, 2018. Ready for malaria elimination: zero indigenous case reported in the People's Republic of China. Malar J 17:315.

Garros C, Van Bortel W, Trung HD, Coosemans M, Manguin S, 2006. Review of the minimus complex of Anopheles, main malaria vector in Southeast Asia: from taxonomic issues to vector control strategies. Trop Med Int Health 11:102-14.

Gosling RD, Whittaker M, Gueye CS, Fullman N, Baquilod M, Kusriastuti R, Feachem RG, 2012. Malaria elimination gaining ground in the Asia Pacific. Malar J 11:346.

Hay SI, Smith DL, Snow RW, 2008. Measuring malaria endemicity from intense to interrupted transmission. Lancet Infect Dis 8:369-78.

Hinrichsen RA, 2009. Population viability analysis for several populations using multivariate state-space models. Ecol Model 220:1197-202.

Holmes EE, Ward EJ, Wills K, 2012. MARSS: Multivariate autoregressive state-space models for analyzing time-series data. R J 4:11-9.

Hu T, Liu YB, Zhang SS, Xia ZG, Zhou SS, Yan J, Cao J, Feng ZC, 2016. Shrinking the malaria map in China: measuring the progress of the National Malaria Elimination Programme. Infect Dis Poverty 5:52.

Koc EK, Bozdogan H, 2015. Model selection in multivariate adaptive regression splines (MARSS) using information complexity as the fitness function. Machine Learn 101:35-58.

Liu H, Huang K, Liu S, Sun H, Lin K, Huang X, Chu J, Yang Z, 2016. Gene frequency and haplotype distribution of hemoglobin E among seven minority groups of Yunnan, China. Am J Hum Biol 28:927-31.

Li BF, Lin YX, Guo XR, Chen LF, Zhou DL, Yu GC, Zou J, Sun XD, 2015. [Survey on malaria epidemics in China-Myanmar border area]. Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 33:261-3 [Article in Chinese].

Li JQ, Yang HX, 2015. [Analysis of time distribution of malaria incidence in Baoshan City during 1990-2014.] Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 33:309-11 [Article in Chinese].

Li S, Yin S, Wang J, Li X, Feng J, 2016. Shifting from control to elimination: analysis of malaria epidemiological characteristics in Tengchong County around China-Myanmar border, 2005-2014. Malar J 15:45.

Ministry of Health, China (MoH), 2010. National Malaria Elimination Action Plan. Beijing, China. Available from:

http://www.nhc.gov.cn/zwgk/wtwj/201304/15a4cc7a40b0452191fe409590ca99d8.shtml

Reiner RC, Le Menach A, Kunene S, Ntshalintshali N, Hsiang MS, Perkins TA, Greenhouse B, Tatem AJ, Cohen JM, Smith DL, 2015. Mapping residual transmission for malaria elimination. Elife. 4:e09520.

Royston JP, 1982. An extension of Shapiro and Wilk's W test for normality to large samples. J Roy Stat Soc Series C (Appl Stat) 31:115-24.

Shi B, Liu J, Zhou XN, Yang GJ, 2014. Inferring Plasmodium vivax transmission networks from tempo-spatial surveillance data. PLoS Negl Trop Dis 8:e2682.

Shi B, Tan Q, Zhou XN, Liu J, 2015. Mining geographic variations of Plasmodium vivax for active surveillance: a case study in China. Malar J 14:216.

Shi B, Zheng J, Qiu H, Yang GJ, Xia S, Zhou XN, 2017. Risk assessment of malaria transmission at the border area of China and Myanmar. Infect Dis Poverty 6:108.

Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz Soldan V, Kochel TJ, Kitron U, Elder JP, Scott TW, 2009. The role of human movement in the transmission of vector-borne pathogens. PLoS Negl Trop Dis 3:e481.

Tolimieri N, Holmes EE, Williams GD, Pacunski R, Lowry D, 2017. Population assessment using multivariate time-series analysis: A case study of rockfishes in Puget Sound. Ecol Evol 7:2846-60.

Valle D, Clark J, 2013. Improving the modeling of disease data from the government surveillance system: a case study on malaria in the Brazilian Amazon. PLoS Comput Biol 9:e1003312.

Wang D, He F, Maslov S, Gerstein M, 2016a. DREISS: using state-space models to infer the dynamics of gene expression driven by external and internal regulatory networks. PLoS Comput Biol 12:e1005146.

Wang D, Li S, Cheng Z, Xiao N, Cotter C, Hwang J, Li X, Yin S, Wang J, Bai L, Zheng Z, Wang S, 2015a. Transmission risk from imported plasmodium vivax malaria in the China-Myanmar border region. Emerg Infect Dis 21:1861-4.

Wang X, Yang L, Jiang T, Zhang B, Wang S, Wu X, Wang T, Li Y, Liu M, Peng Q, Zhang W, 2016b. Effects of a malaria elimination program: a retrospective study of 623 cases from 2008 to 2013 in a Chinese county hospital near the China—Myanmar border. Emerg Microbes Infect 5:e6.

Wang JZ, Yin SQ, Li SG, Li XS, Cai WB, Feng J, 2016c. [Evaluation of measures and achievements of malaria control in Tengchong City, Yunnan Province during 2010-2015.] Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 34:435-8 [Article in Chinese].

Wang RB, Zhang QF, Zheng B, Xia ZG, Zhou SS, Tang LH, Gao Q, Wang LY, Wang RR, 2014. Transition from control to elimination: impact of the 10-year global fund project on malaria control and elimination in China. Adv Parasitol 86:289-318.

Wang Y, Zhong D, Cui L, Lee MC, Yang Z, Yan G, Zhou G, 2015b. Population dynamics and community structure of Anopheles mosquitoes along the China-Myanmar border. Parasit Vectors 8:445.

Wangdi K, Singhasivanon P, Silawan T, Lawpoolsri S, White NJ, Kaewkungwal J, 2010. Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan. Malar J 9:251.

Ward EJ, Chirakkal H, Gonzalezsuarez M, Auriolesgamboa D, Holmes EE, Gerber LR, 2010. Inferring spatial structure from time-series data: using multivariate state-space models to detect metapopulation structure of California sea lions in the Gulf of California, Mexico. J Appl Ecol 47:47-56.

Whittaker MA, Dean AJ, Chancellor A, 2014. Advocating for malaria elimination - learning from the successes of other infectious disease elimination programmes. Malar J 13:221.

World Health Organization (WHO), 2019. Malaria. Available from: http://www.who.int/mediacentre/factsheets/fs094/en/

Xu JW, Li JJ, Guo HP, Pu SW, Li SM, Wang RH, Liu H, Wang WJ, 2017. Malaria from hyperendemicity to elimination in Hekou County on China-Vietnam border: an ecological study. Malar J 16:66.

Xu JW, Li Y, Yang HL, Zhang J, Zhang ZX, Yang YM, Zhou HN, Havumaki J, Li HX, Liu H, Zhou H, Xie XY, Dong JX, Zhang Y, Sun XY, Li B, Li JY, Tian YH, Wang PY, Li BF, 2016. Malaria control along China-Myanmar Border during 2007-2013: an integrated impact evaluation. Infect Dis Poverty 5:75.

Xu JW, Liu H, 2016. The relationship of malaria between Chinese side and Myanmar's five special regions along China-Myanmar border: a linear regression analysis. Malar J 15:368.

Yin JH, Zhou SS, Xia ZG, Wang RB, Qian YJ, Yang WZ, Zhou XN, 2014. Historical patterns of malaria transmission in China. Adv Parasitol 86:1-19.

Yu G, Yan G, Zhang N, Zhong D, Wang Y, He Z, Yan Z, Fu W, Yang F, Chen B, 2013. The Anopheles community and the role of Anopheles minimus on malaria transmission on the China-Myanmar border. Parasit Vectors 6:264.

Zhang S, Guo S, Feng X, Afelt A, Frutos R, Zhou S, 2017. Manguin S. Anopheles vectors in mainland China while approaching malaria elimination. Trends Parasitol 33:889-900.

Zhang Q, Sun J, Zhang Z, Geng Q, Lai S, Hu W, Clements AC, Li Z, 2016. Risk assessment of malaria in land border regions of China in the context of malaria elimination. Malar J 15:546.

Zhang SS, Zhou SS, Zhou ZB, Chen TM, Wang XZ, Shi WQ, Jiang WK, Li JL, Zhou XN, Frutos R, Manguin S, Afelt A, 2018. Monitoring of malaria vectors at the China-Myanmar border while approaching malaria elimination. Parasit Vectors 11:511.

Zhou P, Wang D, Hwang J, Bennett A, Sturrock HJ, Ma A, Huang J, Xia Z, Feng X, Wang J, 2015a. Spatial-temporal variation and primary ecological drivers of Anopheles sinensis human biting rates in malaria epidemic-prone regions of China. PLoS One 10:e0116932.

Zhou SS, Wang Y, Li Y, 2011. [Malaria situation in the People's Republic of China in 2010.] Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi 29:401-3. [Article in Chinese].

Zhou SS, Zhang SS, Zhang L, Rietveld AE, Ramsay AR, Zachariah R, Bissell K, Van den Bergh R, Xia ZG, Zhou XN, Cibulskis RE, 2015b. China's 1-3-7 surveillance and response strategy for malaria elimination: Is case reporting, investigation and foci response happening according to plan? Infect Dis Poverty 4:55.