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,¹⁴ Benyun Shi,⁵ Shang Xia,¹⁴ Guojing Yang,⁶,⁷,⁸ Xiao-Nong Zhou¹⁴

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

Abstract

The transition from the control phase to elimination of malaria in China through the national malaria elimination programme has focused 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.

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 patterns 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 km² 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.](image-url)
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 inaccuracies. The MARSS framework is given by two equations:

\[
X_t = B X_{t-1} + u_t + w_t, w_t \sim N(0, Q)
\]

\[
Y_t = Z X_t + a_t + v_t, v_t \sim N(0, R)
\]

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 \( t \); \( B \), 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 \( Q \). In Eq (2), \( Y_t \) is the \( n \) element vector of observed malaria cases at time \( t \), \( Z \), 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 states. 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 \( R \). 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:

\[
x_t = x_{t-1} + u_t + w_t, w_t \sim N(0, Q)
\]

\[
y_t = 1 \times x_t + a_t + v_t, v_t \sim MVN(0, R)
\]

\[
y_t = 1 \times x_t + a_t + v_t, v_t \sim MVN(0, R)
\]

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 confidence 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 occurred 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 larger 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.
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 reported 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.

<table>
<thead>
<tr>
<th>Model</th>
<th>Q^2</th>
<th>R^2</th>
<th>U^2</th>
<th>LogLik</th>
<th>AI(C)^</th>
<th>AIC^</th>
<th>Parameters^</th>
<th>M^</th>
<th>Hypothesis^</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-121.8388</td>
<td>286.4176</td>
<td>281.6676</td>
<td>19</td>
<td>3</td>
<td>T-S-LYC</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-122.4028</td>
<td>287.5555</td>
<td>282.8055</td>
<td>19</td>
<td>3</td>
<td>S-C-TLY</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-122.7177</td>
<td>288.1853</td>
<td>283.4353</td>
<td>19</td>
<td>3</td>
<td>T-LS-YC</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-122.8298</td>
<td>288.4095</td>
<td>283.6995</td>
<td>19</td>
<td>3</td>
<td>S-TL-YC</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-118.3369</td>
<td>292.4157</td>
<td>284.6738</td>
<td>24</td>
<td>4</td>
<td>T-L-S-YC</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-127.4923</td>
<td>287.9114</td>
<td>284.8464</td>
<td>15</td>
<td>2</td>
<td>SC-TLY</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-127.5118</td>
<td>287.9504</td>
<td>285.0236</td>
<td>15</td>
<td>2</td>
<td>T-LYS</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-127.5694</td>
<td>288.0657</td>
<td>285.1389</td>
<td>15</td>
<td>2</td>
<td>S-TLYC</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-123.9823</td>
<td>290.7146</td>
<td>285.9646</td>
<td>19</td>
<td>3</td>
<td>T-LY-SC</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-118.9946</td>
<td>293.7311</td>
<td>285.8981</td>
<td>24</td>
<td>4</td>
<td>L-S-C-TY</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-124.0113</td>
<td>290.7225</td>
<td>286.0225</td>
<td>19</td>
<td>3</td>
<td>L-S-TYC</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-124.1037</td>
<td>290.9574</td>
<td>286.2074</td>
<td>19</td>
<td>3</td>
<td>C-TY-LS</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-124.2683</td>
<td>291.2885</td>
<td>286.5385</td>
<td>19</td>
<td>3</td>
<td>T-C-LYS</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>-128.523</td>
<td>290.0815</td>
<td>287.1047</td>
<td>15</td>
<td>3</td>
<td>S-C-TLY</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-128.6253</td>
<td>290.2313</td>
<td>287.3045</td>
<td>15</td>
<td>2</td>
<td>TV-LSC</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-124.8939</td>
<td>292.5287</td>
<td>287.7787</td>
<td>19</td>
<td>3</td>
<td>S-TY-LC</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-129.8859</td>
<td>290.3244</td>
<td>287.779</td>
<td>14</td>
<td>2</td>
<td>LS-TYC</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-128.9884</td>
<td>290.9005</td>
<td>287.9373</td>
<td>15</td>
<td>2</td>
<td>TL-YSC</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-125.0074</td>
<td>292.7648</td>
<td>288.0148</td>
<td>19</td>
<td>3</td>
<td>L-TY-SC</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-129.2128</td>
<td>291.3511</td>
<td>288.4243</td>
<td>15</td>
<td>3</td>
<td>S-TC-LY</td>
</tr>
</tbody>
</table>

\[^{1}\text{Process errors with 1 for independent variances and 2 for the same variance; }^{2}\text{observation errors with 1 for independent variances and 2 for same variance; }^{3}\text{reduction rate with 1 for inequality among the data; }^{4}\text{Log-likelihood of the data conditioned on the model parameters; }^{5}\text{Corrected Akaike information criterion; }^{6}\text{Akaike information criterion; }^{7}\text{Number of parameters in the model; }^{8}\text{Hypothesis of hidden states among the five sites; }^{9}\text{Counties grouped by hidden states. T, Tengchong; L, Longling; Y, Longyang; S, Shidian; C, Changning.}\]

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

<table>
<thead>
<tr>
<th>Cluster</th>
<th>County</th>
<th>Reduction rate</th>
<th>Observation error</th>
<th>Process error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tengchong</td>
<td>-0.0223 (-0.1783; 0.4858)</td>
<td>0.0135 (0.0048; 0.0910)</td>
<td>0.2272 (0.0771; 0.3261)</td>
</tr>
<tr>
<td>2</td>
<td>Shidian</td>
<td>-0.0537 (-0.0909; 0.0165)</td>
<td>0.1517 (0.0884; 0.2320)</td>
<td>0.0509 (0.0003; 0.1010)</td>
</tr>
<tr>
<td>3</td>
<td>Longling</td>
<td>-0.0283 (-0.0240; 0.0753)</td>
<td>0.2229 (0.0570; 0.1298)</td>
<td>0.0889 (0.0218; 0.1455)</td>
</tr>
<tr>
<td></td>
<td>Longyang</td>
<td>0.2074 (0.1248; 0.3108)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Changning</td>
<td>0.3125 (0.1771; 0.4858)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 \textit{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 vice 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 respect 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 mosquito 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 situation 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 \textit{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 simulation 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 socioeconomic factors 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 identify 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 which should help supporting long-term policies.

**References**


http://www.nhc.gov.cn/zwgk/wtjj/201304/15a4cc7a40b0452191e40950caf998.shtml


