

Original Paper

Cite this article: Chadsuthi S, Althouse BM, Iamsirithaworn S, Triampo W, Grantz KH, Cummings DAT (2018). Travel distance and human movement predict paths of emergence and spatial spread of chikungunya in Thailand. *Epidemiology and Infection* **146**, 1654–1662. <https://doi.org/10.1017/S0950268818001917>

Received: 19 September 2017
Revised: 21 May 2018
Accepted: 13 June 2018
First published online: 9 July 2018

Key words:

Chikungunya virus; gravity model; human movement; spatial spread; Thailand

Author for correspondence:

S. Chadsuthi,
E-mail: schadsuthi@gmail.com,
sudarac@nu.ac.th

Travel distance and human movement predict paths of emergence and spatial spread of chikungunya in Thailand

S. Chadsuthi¹, B. M. Althouse^{2,3,4}, S. Iamsirithaworn⁵, W. Triampo^{6,7,8},
K. H. Grantz^{9,10} and D. A. T. Cummings^{9,10,11}

¹Department of Physics, Research Center for Academic Excellence in Applied Physics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand; ²Institute for Disease Modeling, Bellevue, WA 98005, USA; ³Information School, University of Washington, Seattle, WA 98105, USA; ⁴Department of Biology, New Mexico State University, Las Cruces, New Mexico 88003, USA; ⁵Department of Disease Control, Ministry of Public Health, Tivanond 9 Road, Nonthaburi 11000, Thailand; ⁶Biophysics Group, Department of Physics, Faculty of Science, Mahidol University, Rama VI, Bangkok 10400, Thailand; ⁷Centre of Excellence in Mathematics CHE, Sriyudhaya Rd., Bangkok 10400, Thailand; ⁸TheEP Center, CHE, 328 Si Ayutthaya Road, Bangkok 10400, Thailand; ⁹Department of Biology, University of Florida, Gainesville, FL 32611, USA; ¹⁰Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611, USA and ¹¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

Abstract

Human movement contributes to the probability that pathogens will be introduced to new geographic locations. Here we investigate the impact of human movement on the spatial spread of Chikungunya virus (CHIKV) in Southern Thailand during a recent re-emergence. We hypothesised that human movement, population density, the presence of habitat conducive to vectors, rainfall and temperature affect the transmission of CHIKV and the spatio-temporal pattern of cases seen during the emergence. We fit metapopulation transmission models to CHIKV incidence data. The dates at which incidence in each of 151 districts in Southern Thailand exceeded specified thresholds were the target of model fits. We confronted multiple alternative models to determine which factors were most influential in the spatial spread. We considered multiple measures of spatial distance between districts and adjacency networks and also looked for evidence of long-distance translocation (LDT) events. The best fit model included driving-distance between districts, human movement, rubber plantation area and three LDT events. This work has important implications for predicting the spatial spread and targeting resources for control in future CHIKV emergences. Our modelling framework could also be adapted to other disease systems where population mobility may drive the spatial advance of outbreaks.

Introduction

Infectious diseases emerge due to a variety of factors. The transmissibility of pathogens, environmental changes, economic development and human behaviour all contribute to the probability of emergence [1]. Human mobility intuitively plays a key role in emergence [2] and can determine the speed at which pathogens are introduced to areas where previously absent. Rvachev and Longini showed the importance of air travel in dictating the timing of emergence of H3N2 influenza in the pandemic of 1968 using a metapopulation model [3]. Recently, networks of human movement have been shown to be integral in predicting and understanding the global spread of influenza A (H1N1), Ebola and SARS [4–6]. Consequently, restriction of travel has been discussed as a potential disease control strategy [2], though the feasibility of these controls may be limited [7].

Rarely, though, has empirical data been used to compare multiple competing transmission models to determine which one shows the greatest correspondence with the spatio-temporal pattern of cases in an outbreak. Here, we demonstrate the importance of human movement in the spatial spread and emergence of a vector-borne disease. We compared models that incorporate different measures of human mobility, different distance metrics between populations (i.e. driving distance vs. geodesic) as well as environmental conditions of each location.

Chikungunya (CHIK) is an acute illness caused by an arthropod-borne alphavirus, Chikungunya virus (CHIKV). CHIKV is transmitted to hosts via bites from an infected mosquito, predominantly of the genus *Aedes* [8, 9]. Chikungunya, translated from Makonde as ‘that which bends up’, is so called due to the severe joint pain in both the acute and chronic stages [10]. The first recognised outbreak of CHIKV occurred in 1952–1953 in present-day Tanzania [11]. The pathogen has since spread widely throughout India and several countries in Southeast Asia [12]. Scattered importations and autochthonous cases have been reported in

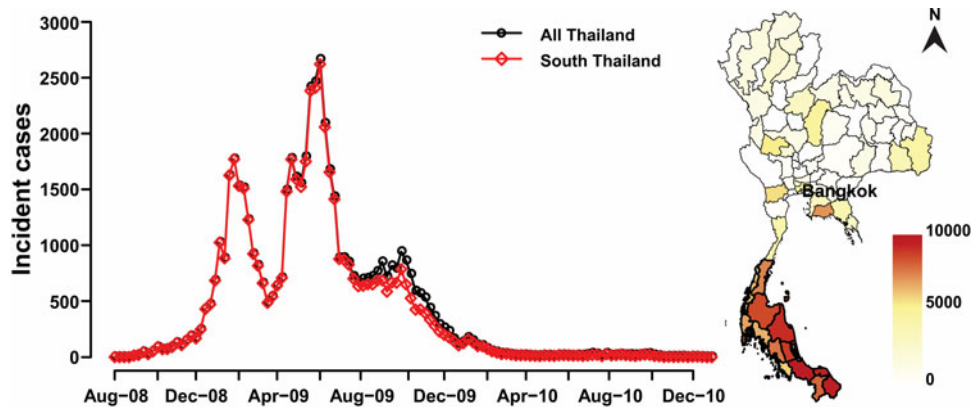


Fig. 1. Time series of the incident reported cases by week (left) and heat map of total reported cases in each province (right). The Southern region studied in this work is outlined.

Europe [13], the USA [14] and sub-Saharan and West Africa [15]. Over one million cases have been reported in the Americas since CHIKV was detected in December 2013 [16].

The first urban outbreaks in Bangkok, Thailand, occurred in the early 1960s [17]. After a 13-year absence from Thailand, an outbreak of CHIKV was reported in 2008 in the province of Narathiwat near the Malaysian border [18]. By the end of 2010, CHIKV had been identified in over one-third of districts in Thailand. Whereas previously circulating strains in Thailand were of Asian lineage [19], the 2008 outbreak was the result of an introduction of the East Central and South African lineage [20, 21]. Sequencing of CHIKV isolated during the 2008 outbreak revealed the presence of a point mutation, A226V, shown to increase vector specificity for *Aedes albopictus* [22].

Several factors have been hypothesised to have influenced the spread of CHIKV in Thailand. We investigate here several weather and environmental covariates that have previously been identified as potentially affecting the transmission of CHIKV. Increased rainfall [23], higher temperatures [24, 25] and increased forest [26] or marshy rubber plantation coverage [27] provide favourable conditions for vector replication and survival and therefore could impact the transmission of CHIKV. Rainfall and ambient temperature may impact the abundance of *Ae. albopictus* [28] through multiple mechanisms, including increased availability of breeding sites [23], increased emergence rates [23], improved larval survival [24], increased biting rates and reductions in the extrinsic incubation period [24, 25]. Human movement and population densities could facilitate geographic spread through human hosts [26]. Long distance movement of infected individuals into uninfected areas, here termed long-distance translocation (LDT) events [29], may have also contributed to the re-emergence and spread of CHIKV. In this work, we used district-level CHIK incidence data to fit metapopulation transmission models to elucidate which of these factors were major determinants of spatial spread. We identified the model for which the data had the maximum likelihood and used likelihood ratio tests to compare models.

Materials and methods

Data

Clinical cases of CHIK in Thailand were reported to the National Surveillance System administered by the Bureau of Epidemiology,

Department of Disease Control in the Ministry of Public Health. CHIKV was first isolated at Yi-ngo, Narathiwat, Thailand in August 2008 [18]. From August 2008 to December 2010, there were a total of 56 112 cases and at least one case of CHIK was detected in 316 of 926 districts in Thailand (see supplemental animation) [30]. Serum samples were obtained from 3434 of 56 112 (6.1%) reported cases for either virologic or serologic testing. Of those samples tested, 1219 (35.5%) were subsequently confirmed as CHIKV infections. The epidemic was focused in Southern Thailand, where 94.8% of all cases occurred. Figure 1 shows incident cases per week in the southern districts and for the entire country. In this work, we focused on districts in the Southern region. A district was considered infected once total incidence exceeded one reported case per 10 000 inhabitants. A total of 127 of 151 southern districts meet this definition during the outbreak. We modelled the process by which districts moved between uninfected and infected states.

We obtained data describing population sizes, rainfall, temperature and percentage forest and rubber plantation district coverage in each of the 151 southern districts. Population sizes were obtained from the Department of Provincial Administration, Ministry of Interior of the Kingdom of Thailand [31].

The fraction forest coverage was calculated as the total area covered by forest, based on Landsat remote sensing [32], divided by the total district area. We considered only forestation, excluding plantations, eucalyptus plantation, secondary forest and other agricultural area. The fraction rubber plantation coverage was calculated in a similar manner using data from the Provincial Agricultural Extension, Department of Agriculture Extension, Thailand [33, 34].

Rainfall data were obtained from the real-time TRMM Multi-Satellite Precipitation Analysis (TMPA-RT) [35]. The daily accumulated precipitation in millimeters was obtained from TRMM 3B42RT Daily for the centroid of each district [36, 37]; we assumed homogeneity of district-level rainfall from that reported at the centroid.

The average temperature was obtained from the Moderate Resolution Imaging Spectroradiometer of the Terra satellite, which records the average value of daytime and nighttime land surface temperature on a 1 km² sinusoidal grid [38]. The average temperature in each district is taken to be the average temperature of all 1 km² image pixels within the district.

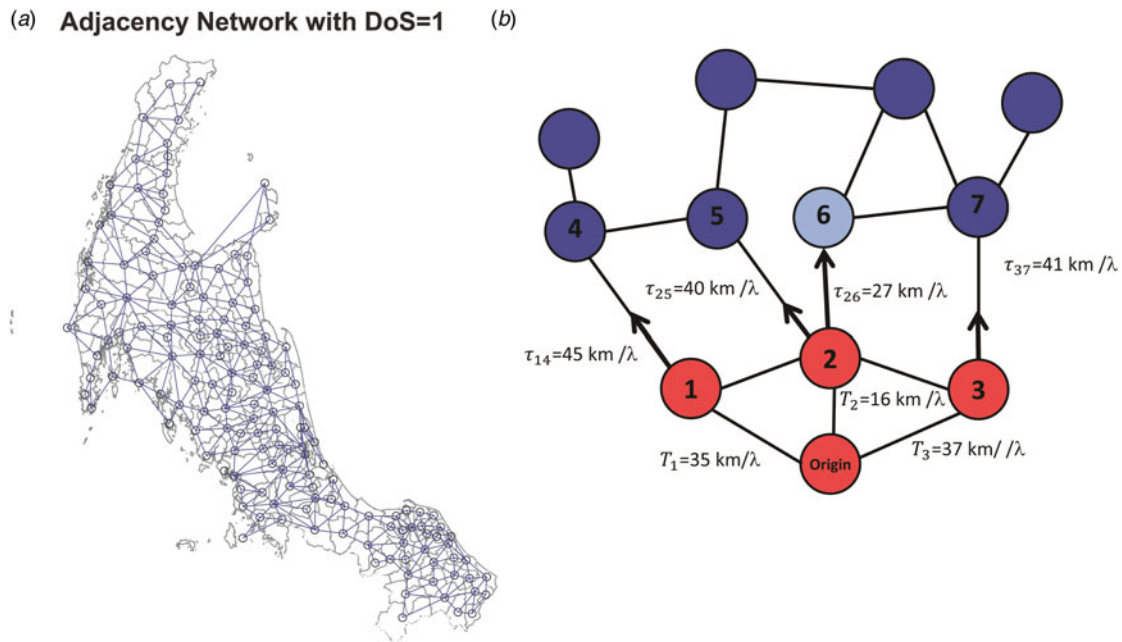


Fig. 2. (a) The network model overlaid on a map of Southern Thailand with DoS = 1. (b) An example of the metapopulation transmission model; red circles represent infected districts and blue circles uninfected districts. For all infected districts at this time point (Districts 1–3), the spatial spread of the infection proceeds along the link with the shortest time to next infection, $T_j = T_k + \tau_{kj}$, (from District 2 to District 6). The process is repeated iteratively until all districts are infected.

Model

Metapopulation transmission model

We built a metapopulation model of transmission across the 151 districts of Southern Thailand to model the spatiotemporal process by which districts became infected as a function of the state of neighbouring districts and network distances. We extended a model developed by Smith *et al.* for examining the dynamics of rabies in raccoons in Connecticut [29]. We assumed that, once a patch reaches a particular threshold, the risk associated with dispersion of cases to other areas is independent of the overall incidence within a patch [39, 40].

Each district was represented by a node in a weighted graph, with each pair of adjacent districts connected by a link (Fig. 2, panel a). We define T_j to be the simulated time in weeks elapsed from the first observed case in Southern Thailand to the time that the j th district reached the infection threshold. T_j was computed as a function of the adjacency network, N , the rate of spatial spread between infected district k and uninfected district j in units of kilometers per week, λ_{kj} and the set of pairwise distances between districts k and j , d_{kj} . Therefore, time of infection in district j , T_j , can be defined as $T_j = T_k + \tau_{kj}$, where τ_{kj} is the expected time for cases to be introduced to district j directly from district k , defined as d_{kj}/λ_{kj} . The weight of each node is assigned to be τ_{kj} .

The times of infection for each district were computed by iteratively identifying the next district to become infected. This was done by identifying at each time point the district, j , which minimised $T_j = T_k + \tau_{kj}$ across all infected districts k . In identifying the sequence of these individual infection events, the spatial progression of districts exceeding the specific threshold is determined for each model parameterisation. For example (Fig. 2, panel b), there may be three currently infected districts from which CHIKV can be spread to any of the four uninfected districts at a constant rate, λ . Possible routes of transmission are

defined by the adjacency network, N . At this time point, the time of infection T_j is least for District 6 ($T_6 = (16 \text{ km} + 27 \text{ km})/\lambda$), so it is said to be newly infected by District 2. The algorithm was repeated until all districts became infected. The output T_j was compared with the observed data to find the optimised model.

We consider three metrics of distance, d_{kj} . First, we measure the geodesic distance ('as the bird flies') from the centroid of each district. We also consider both the geodesic and driving distances (as measured by Google Maps [41]) between the administrative offices of each district, which are typically located in areas with the largest population density in each district and may more accurately reflect travel distances between districts.

We also define several different adjacency networks to explore whether the data were most consistent with spread only between neighbouring districts, or with spread occurring from neighbours of neighbours or beyond. Adjacency matrices defined the degree of separation (DoS) between two districts, an integer number that defined the minimum number of shared boundaries that must be crossed to reach one district from another. If two districts were adjacent (i.e. share a boundary), they were considered to be separated by 1° of separation (DoS = 1). In this work, we considered a model that included transmission events between districts with a DoS of 2 or lower. However, this model does not perform as well as DoS = 1.

For districts that remained uninfected for the length of the observed data, we arbitrarily set the time of infection to the week after the latest observed data point. This reflects the fact that our observations of time of infection in each district is censored. We explored the assumption of the arrival time for these uninfected districts and found that model results were insensitive to the specific assumption (e.g., 1 week or 2 weeks after last observation) of the arrival time for these districts. These putatively uninfected districts can infect and spread to other districts in the model.

We tested multiple forms of the rate of spread between districts, λ_{kj} , as a function of heterogeneous environmental variables to examine the influence of various factors on CHIKV transmission. We used a simple rule for generating the rates of spread. We examined five forms for λ_{kj} corresponding to five environmental variables:

- (1) Homogeneous, $\lambda(H_0)$: constant rate of spread, i.e. $\lambda_{kj} = \alpha$ for all k and j .
- (2) Human movement (gravity model), $\lambda(H)$: here, λ_{kj} depends on the population density of nodes k and j , through a gravity model, where the probability that an individual visits a district is directly proportional to the population sizes of the origin (N_k) and destination districts (N_j) [42, 43]. We took $\lambda_{kj} = \max(\alpha(1 + \theta\{(N_j^{p_1} N_k^{p_2})/d_{kj}^\sigma\}), 0)$, where θ is a constant, p_1 , p_2 and σ are the exponents and d_{kj} is the geodesic or driving distance between districts k and j .
- (3) Weather conditions, $\lambda(C)$: we assumed the rate of spread depends on rainfall or temperature conditions at the time step before infection (T_k) in the infecting and infected districts. Weather variation could affect the transmissibility of CHIKV by influencing the mosquito vector life cycle or growth of the virus in the vector.
 - (3.1) Rainfall, $\lambda(R)$: $\lambda_{kj} = \max(\alpha(1 + \rho(\text{Rain}_k(T_k) + \text{Rain}_j(T_k))), 0)$, where Rain_j is rainfall in mm for the district j and ρ is a constant.
 - (3.2) Temperature, $\lambda(T)$: $\lambda_{kj} = \max(\alpha(1 + \gamma(\text{Temp}_k(T_k) + \text{Temp}_j(T_k))), 0)$, where Temp_j is the temperature in degrees C for the district j and γ is a constant.
- (4) Forest, $\lambda(F)$ and rubber plantation, $\lambda(P)$: we assumed the rate of spread is linearly proportional to the percent of the forest, F_j , or rubber plantation, P_j , an area in the j th district: $\lambda_{kj} = \max(\alpha(1 + \beta F_j), 0)$ or $\lambda_{kj} = \max(\alpha(1 + \beta P_j), 0)$. The forest or rubber plantation areas provide suitable breeding habitats for CHIKV vectors [26, 27].

We also considered linear combinations of the above rates in multi-factor models. For example, $\lambda(H, R, F)$ is the combination model of human movement, rainfall and the percent of the forest, which equal to $\alpha(1 + \theta\{(N_j^{p_1} N_k^{p_2})/d_{kj}^\sigma\} + \rho(\text{Rain}_k(T_k) + \text{Rain}_j(T_k)) + \beta F_j)$, while $\lambda(H, R, T, F)$ is the combination model of human movement, rainfall, temperature and the percent of the forest, which is equal to $\alpha(1 + \theta\{(N_j^{p_1} N_k^{p_2})/d_{kj}^\sigma\} + \rho(\text{Rain}_k(T_k) + \text{Rain}_j(T_k)) + \gamma(\text{Temp}_k(T_k) + \text{Temp}_j(T_k)) + \beta F_j)$. The weight of each rate of spread was adjusted by a constant α so that, if one of the environmental coefficients was equal to zero, the transmission rate collapsed to the homogeneous model.

LDT events

Epidemics may not diffuse contiguously between neighbouring districts due to infectious individuals travelling beyond neighbouring districts. LDT events could progress the spread of advancing epidemics across multiple boundaries at speeds independent of the rates described above. We identified potential LDT events and included them in the model to improve fit. In constructing candidate models, our criteria for identifying LDT events were:

- LDTs must be at least 2 DoS from the most recently infected districts,
- The LDT must be the first reported case in the focal district and its neighbouring districts (DoS = 1),

- The LDT must be before the median week of cases in the province containing the focal district (Supplemental Figure S1 shows province and district boundaries).

Estimation

To find the best model, we simulated all possible combinations of λ , including multi-factor models, for each of the three possible pairwise distance metrics. We compared models using the maximum likelihood estimate. The best fit was found by maximising a normal log-likelihood, which found the simulation results that were most likely from the observed data [29]. We used a likelihood of the form:

$$L(\mu, \sigma, x) = \sigma^{-n} (2\pi)^{-(n/2)} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right]$$

where the simulated arrival date appears as the expectation of the random variable and the observed date as x_i .

We used Nelder–Mead optimisation to set initial parameter values, which is the simplex method to minimise a function of n variables [44]. Then, we used a quasi-Newton method (Broyden–Fletcher–Goldfarb–Shannon, BFGS) to choose new search directions. This technique searches the solution in the vicinity based on mathematical gradients, which dictate the convergence rate of response surface methodology [45, 46]. The likelihood ratio test was used to calculate the 95% partial likelihood-based confidence interval of fitted parameters.

Results

Twenty-four models were fitted to the observed data for each distance metric (Fig. 3). In general, we found that human movement and rainfall improved fit when included as single factors, while rubber plantation models fit the data poorly. The optimal homogeneous model, $\lambda(H_0)$, gives an average rate of spread of 8.5 km/week ($R^2 = 0.9868$), indicating that CHIKV spread throughout Southern Thailand from the outbreak source in 82.3 weeks, slower than the observed 61 weeks between the first and last district infection events.

The human movement model using straight line distance between district offices was the best one-factor model, with a reduction in log likelihood of 21 332.12 from the homogeneous transmission model. This result highlights the importance of human movement in CHIKV spread (Fig. 3). Rubber plantation area alone may not be significant to the transmission dynamics of CHIKV, as these models showed the worst performance of all the one-factor models. The inclusion of weather conditions in the formulation of $\lambda(R, T)$ slightly improved fit, with rainfall and temperature having similar effects when compared with one-factor model of rainfall or temperature.

The HP (human movement and rubber plantation coverage) model had the lowest negative log-likelihood among models with driving distance. The HF (human movement and forest coverage) model showed the best performance in straight-line office distance, while the RP (rainfall and rubber plantation coverage) model showed the best performance in straight-line centroid distance. The four-factor HRTF model fits better than the HRTF model using any of the three-distance metrics, though neither performed as well as the HP model with driving distance (Fig. 3). Thus, we selected the HP using driving distance model,

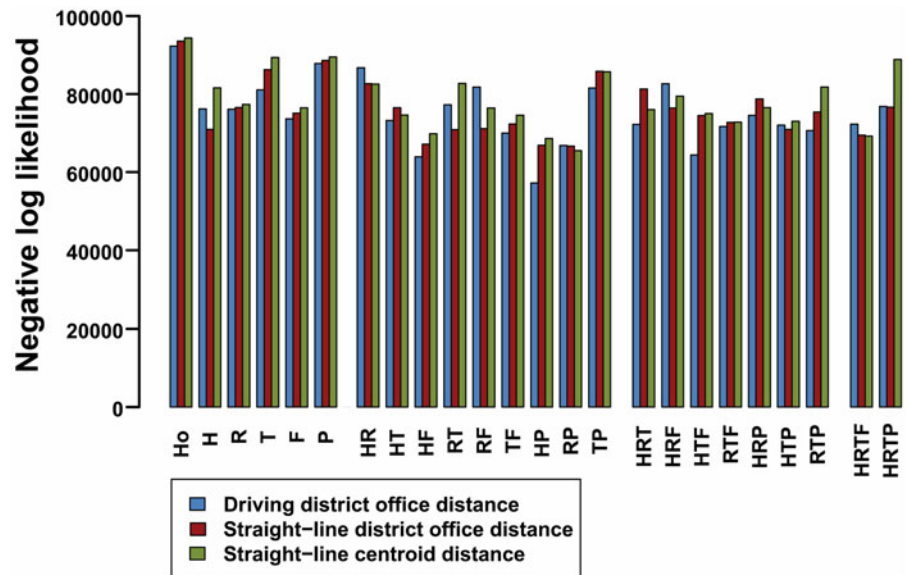


Fig. 3. Negative log likelihood for the 24 candidate models of CHIKV spread with three different metrics of between-district distance. Ho indicates the null model, the homogenous rate of spread between districts. Other models indicate single- and multi-factor models of the rate of spread. H, human movement; R, rainfall; T, temperature; F, forest; and P, rubber plantation.

which has maximum log likelihood (minimum negative log likelihood) and fewer parameters, as the best-fit model for further study.

Maps of residual error for one-factor models using driving distance are presented in Fig. 4. Overall, the one-factor models predicted later emergence of CHIKV than observed in Southern Thailand. The error was distributed across the region, with generally early prediction in the mid-south, near the first infected district and late predictions in the northern and southernmost districts. The rainfall model had the lowest overall residual error among the one-factor models. The multi-factor HP model showed the largest reduction in error compared to the other models, including reduction of late-prediction errors observed in the one-factor models.

LDT events could progress the spread of advancing epidemics across multiple districts at speeds independent of the rates described above and account for the non-contiguous spread of CHIKV. Due to limitations in computational time, LDT events were considered only in the best fit multi-factor model (HP). A total of 22 districts were found to meet the criteria described above for CHIKV introduction through an LDT event. We tested each possible LDT event independently to see if its inclusion improved the model fit and found 14 LDT events that improved model fit by maximum log-likelihood estimator (MLE; Supplemental Table S1). The LDT event that showed the greatest increase in MLE was an introduction into Kathu, Phuket, located in the middle-west part of southern (Supplemental Figure S2). The worst performing LDT event was an introduction into Saba Yoi, Songkhla. We sequentially added additional LDT events, testing all possible sets of 1, 2, 3, 4 and 14 LDT events. The model with 14 LDT events (all events that improved the performance of the model when added independently) was tested as a model with a high degree of freedom.

Model fit observably improved with the addition of more than one LDT event (Fig. 5). We found that the inclusion of all 14 LDT events did not result in the best performance. Rather, the HP model with 3 LDT events, with introductions into Thalang, Phuket; BanTaKhun, Surat Thani; and YanTaKhao, Trang, was selected to be the best fitting models (Fig. 6). The best-fit HP model with 3 LDT resulted in an absolute difference of log

likelihood of 614.17 from the best fitting model including 4 LDT events. The 3 LDT model also demonstrated strong positive correlation (0.74, P -value <0.001 using Pearson correlation test) to the observed data (Fig. 6). The fitted parameters and 95% confidence interval are shown in Supplemental Table S2. Again, the most residual error involved early prediction of CHIKV emergence in southern districts. The most likely network model indicated that CHIKV moved adjacently district-to-district with few LDT events (Fig. 6, panel c).

Discussion

In this work, we presented a metapopulation transmission model of the invasion and spread of Chikungunya virus across districts in Southern Thailand from 2008 to 2010. Of the factors investigated in this study (human movement, rainfall, temperature, forest coverage and rubber plantation coverage), we found human movement and rubber plantation coverage (HP model) most improved fit to the observed data. We also identified three LDT events that statistically significantly improved model fit.

Our analysis demonstrated that driving distance between district offices offers a better fit to the observed data than straight line distance between districts. This result lends credence to the importance of human movement in the spread of CHIKV. While straight line distance may have some influence on human behaviour regarding travel (longer distances being less frequented), the actual movement of individuals is likely dictated more by travel or driving distance. The road infrastructure and landscape of Southern Thailand may limit the frequency of travel, even between geographically close districts [47] and therefore may restrict the paths of CHIKV emergence and movement.

Implicit in the use of a gravity model of human movement is the dependence of movement patterns on population density. Here, metapopulation models dependent on distance alone showed much worse fit than models incorporating a gravity model of human movement. Thus, two key drivers of human movement, travel distance and population density, are important in the movement of CHIKV across Thailand. This echoes previous findings explaining transmission dynamics of seasonal influenza [43] and dengue introductions in the USA [48].

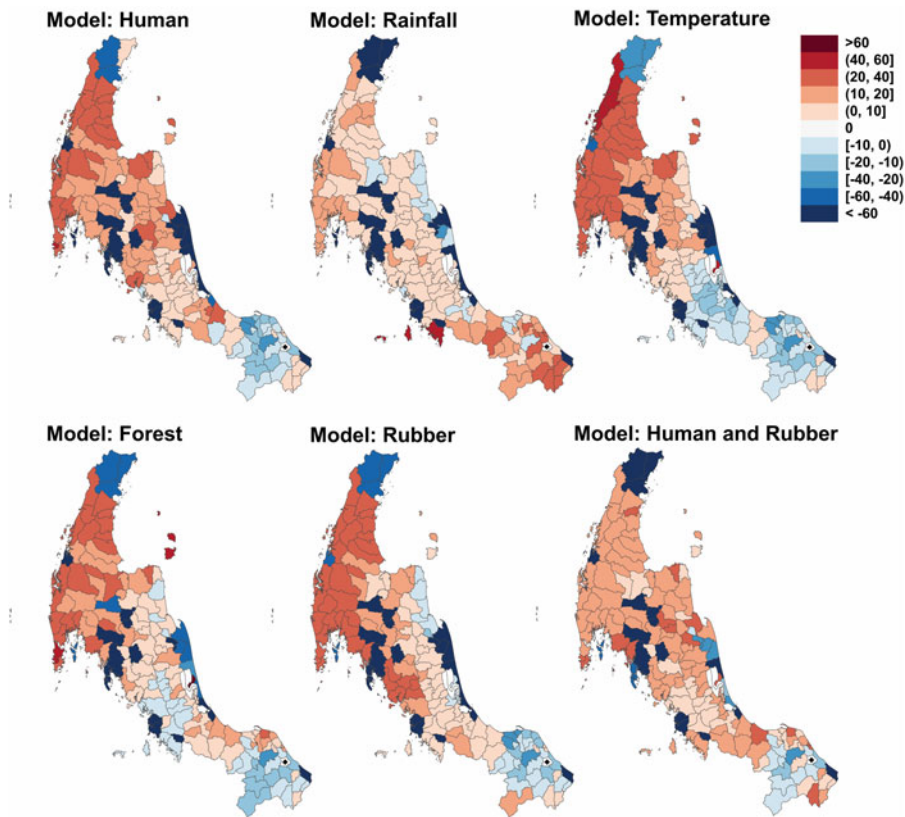


Fig. 4. Comparison of residual errors of driving distance model of the one-factor models and the best-factors (HP) model. Positive errors indicate late prediction; negative errors indicate early prediction. Black dots indicate the first infected district.

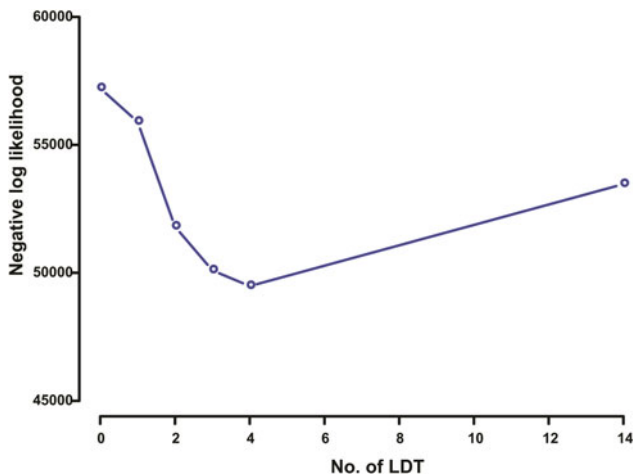


Fig. 5. Negative log likelihood estimates of the best-fit model (HP model) with the several numbers of LDT events.

CHIKV strains isolated during the 2008–2010 outbreak in Thailand were most closely related to African genotypes and to isolates from a 2007 outbreak in India, suggesting the novel introduction of CHIKV to Thailand, rather than descent from previously circulating strains [20]. All isolates contained the A226V mutation, known to enhance viral infectivity and dissemination in *Ae. albopictus* and to have an overall selective advantage in *Ae. albopictus*, but not in *Ae. aegypti* [22]. In this outbreak, blood meals from wild-caught mosquitoes showed a significantly higher infection rate in *Ae. albopictus* than in *Ae. aegypti* [28].

Climate may affect the abundance and life cycle of *Ae. albopictus* [49] and therefore have consequences for chikungunya

transmission. In this work, we included two weather factors (rainfall and temperature) known to influence vector survival and the extrinsic incubation period of CHIKV within the vector [24, 25] in models of CHIKV spread. Although neither rainfall nor temperature was included in the best-fit HP model, they did generally improve model fit, notably in one-factor models. In this model, we considered only district-level average temperatures and precipitation. Recent work has indicated that a single extreme rainfall event could increase the abundance of *Ae. albopictus* and consequently extend CHIKV transmission period [23]. Weather events may also influence human mobility patterns, by modulating travel demand or increasing travel times. Increasing rainfall and higher temperatures create more favourable conditions for the vectors that spread CHIKV and may also influence the extrinsic incubation period of CHIKV within the vector [24, 25]. Rainfall may influence in combination with mobility. On a rainy day, travel demand may decrease. During periods of high rainfall, travel time could increase as drivers tend to reduce their speed, exacerbating distances between places.

Rubber plantation coverage was included in the best fit multifactor model, though we found little improvement in model fit in the one-factor model of rubber plantation coverage. Rubber farming is a major agricultural occupation in Southern Thailand which may drive human migration patterns. The marshy conditions of rubber plantations provide favourable environmental conditions for vector growth and increased human exposure to these vectors [27, 50]. Farmers work under shaded rubber trees, where *Ae. albopictus* is abundant and latex cups used to collect rubber provide could increase the number of available breeding sites [27, 51].

Forested land is also an important habitat for *Ae. albopictus* and therefore may play a role in CHIKV transmission [26, 51],

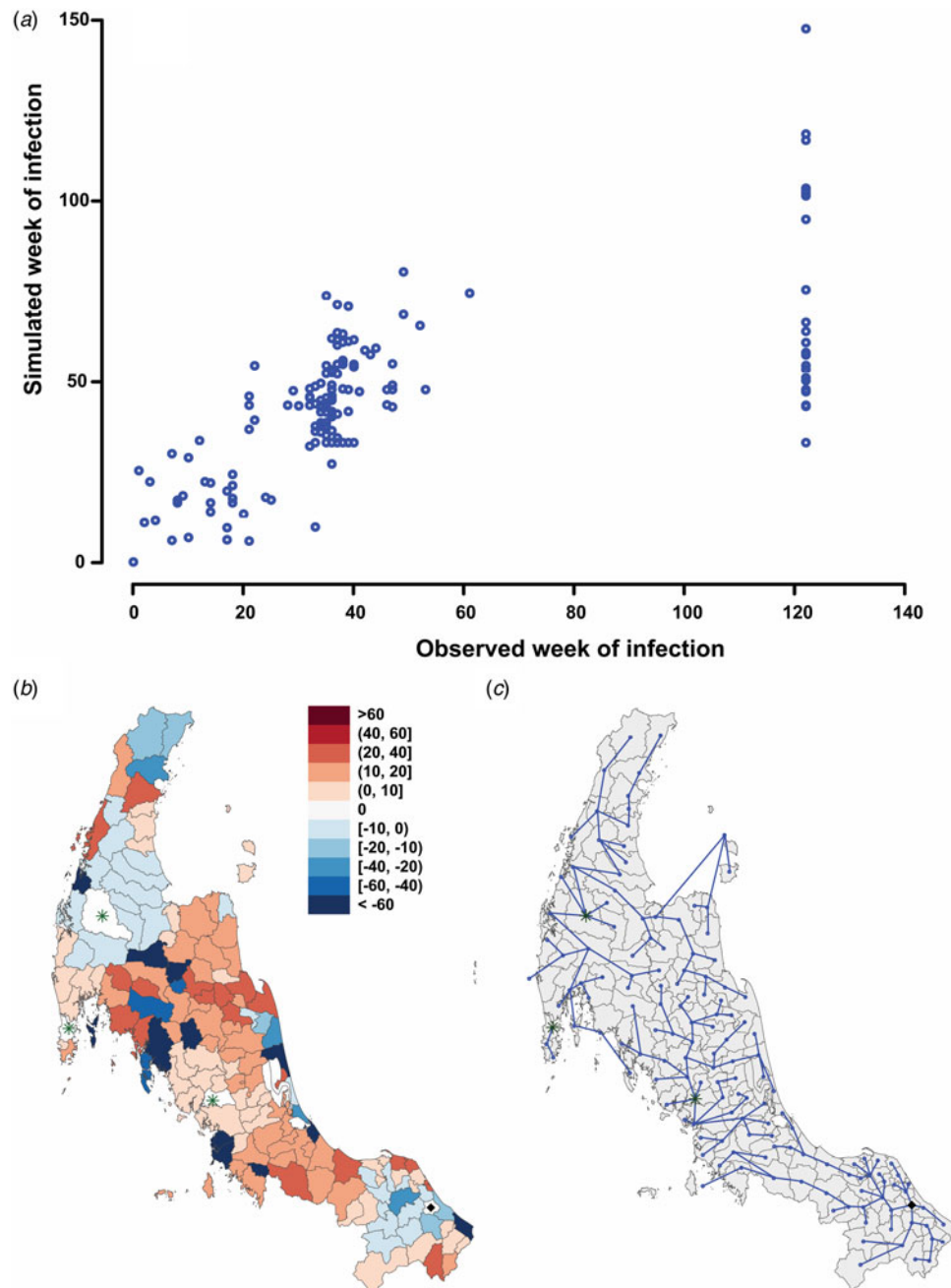


Fig. 6. Scatter plot compares the simulated and observed week of chikungunya from the model fit (a). Maps show the residual error (b) and most likely network model (c) of the best-fit HP model with 3 LDT events (Thalang, Phuket; BanTaKhun, Surat Thani; and YanTaKhao, Trang). Black dots indicate the first outbreak district. Green stars indicate the location of LDTs.

though the impact of forest coverage is likely to depend on land use patterns. If people spend less time in forested areas where *Ae. albopictus* is present, increased forest coverage could be associated with less transmission, particularly if high forest coverage is associated with fewer areas of overlapping vector abundance and human activity, such as rubber plantations. This may explain why forest coverage is not included in the best fit model but does improve model fit in univariate models (Fig. 3). Our results indicate that CHIKV spread throughout Southern Thailand in a relatively linear fashion, with introduction and transmission largely limited to neighbouring districts. Three LDTs substantially improved model fit: BanTaKhun, Surat Thani; Thalang, Phuket;

and YanTaKhao, Trang. BanTaKhun is located about 447 km from the origin of the outbreak (Yi-ngo, Narathiwat) above a geographical bottleneck at about 7.5° of latitude, which separated areas of relatively slow transmission to the south and faster transmission to the north (Fig. 6). The presence of this difference in rates of spread lends evidence to this LDT. The second and third identified LDT events were to Thalang, Phuket and YanTaKhao, Trang, both near the border of the bottleneck.

While including long-distance translocations of CHIKV in Thailand did improve model fit, these events were less critical to understanding transmission dynamics than in the work of Smith *et al.* to understand raccoon rabies in Connecticut [29].

The geography of each region and disease host may help explain this difference. Certain geographic features, such as rivers, are more likely to affect raccoon movement on small spatial scales than human movement. This could limit the ability for populations – and therefore disease – to move between neighbouring regions and increase the importance of LDTs in understanding spatial spread. We employed a standard gravity model to represent human movement, which relies on relatively simplistic assumptions that ignore individual preferences (including influences of social networks) when modelling movement. However, little or no data exist specifically related to human movement in Thailand and the gravity model remains an important approximation of travel patterns. The resolution and accuracy of data describing weather conditions, forest coverage and rubber plantation area are also limited due to the available data.

We assume in multi-factor models that different covariates combine linearly to influence rates of transmission. We do not consider possible interactions or non-linear combinations of multiple factors. Furthermore, we focus on the effect of human movement without directly considering heterogeneity in vector abundance or movement, which are important factors when designing vector control strategies. The accuracy and resolution of district-level weather and land use covariates is limited due to the available data and, due to the nature of CHIKV disease, case reports may be of high specificity [20]. Limitations of the National Surveillance System meant some samples were tested for the presence of CHIKV only by serology, a less sensitive measure than RT-PCR. Some of the observed error in final model fit may be due to asymptomatic transmission or poor reporting to the National Surveillance System.

The significance of the association of human movement in CHIKV movement across Thailand and the observation that CHIKV spread predominantly between neighbouring districts (as demonstrated by statistical support for a small number of LDT) will aid in predicting paths of emergence and preparing for future outbreaks of chikungunya. Entwined with this understanding is the importance of passive surveillance in tracking ongoing outbreaks. This work has clear applications to other disease systems where human movement and physical attributes of locations contribute to the speed of transmission across landscapes.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0950268818001917>.

Acknowledgements. This work was partially supported by Faculty of Science, Naresuan University, Faculty of Science, Mahidol University and the Thailand Center of Excellence in Physics. DATC was supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund. DATC was also supported by the Bill and Melinda Gates Foundation and the US National Institute of General Medical Sciences (Grant 5U54GM088491, Computational Models of Infectious Disease Threats). We thank the Bureau of Epidemiology, Thailand Ministry of Public Health for providing surveillance data.

Declaration of interest. None.

References

1. Wilson ME (1995) Travel and the emergence of infectious diseases. *Emerging Infectious Diseases* **1**, 39–46.
2. Bajardi P *et al.* (2011) Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PLOS ONE* **6**, e16591.
3. Rvachev LA and Longini I (1985) A mathematical model for the global spread of influenza. *Mathematical Biosciences* **75**, 1–22.
4. Fraser C *et al.* (2009) Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* **324**, 1557–1561.
5. Alexander KA *et al.* (2015) What factors might have led to the emergence of Ebola in West Africa? *PLOS Neglected Tropical Diseases* **9**, e0003652.
6. Han Z *et al.* (2014) Effect of human movement on airborne disease transmission in an airplane cabin: study using numerical modeling and quantitative risk analysis. *BMC Infectious Diseases* **14**, 434.
7. Read JM, Eames KT and Edmunds WJ (2008) Dynamic social networks and the implications for the spread of infectious disease. *Journal of The Royal Society Interface* **5**, 1001–1007.
8. Nyari N *et al.* (2016) Identification and genetic characterization of chikungunya virus from Aedes mosquito vector collected in the Lucknow district, North India. *Acta Tropica* **158**, 117–124.
9. Armstrong PM *et al.* (2017) Northern range expansion of the Asian tiger mosquito (*Aedes albopictus*): analysis of mosquito data from Connecticut, USA. *PLOS Neglected Tropical Diseases* **11**, e0005623.
10. Charrel RN, de Lamballerie X and Raoult D (2007) Chikungunya outbreaks – the globalization of vectorborne diseases. *New England Journal of Medicine* **356**, 769–771.
11. Robinson MC (1955) An epidemic of virus disease in Southern Province, Tanganyika territory, in 1952–53. I. Clinical features. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **49**, 28–32.
12. Sudeep AB and Parashar D (2008) Chikungunya: an overview. *Journal of Biosciences* **33**, 443–449.
13. European Centre for Disease Prevention and Control (2018) Stockholm: ECDC. Available at https://ecdc.europa.eu/sites/portal/files/documents/AER_for_2015-chikungunya.pdf.
14. Leparc-Goffart I *et al.* (2014) Chikungunya in the Americas. *The Lancet* **383**, 514.
15. World Health Organization (2015) Available at <http://www.who.int/media-centre/factsheets/fs327/en/>. Accessed May 2015.
16. Fischer M and Staples JE (2014) Notes from the field: chikungunya virus spreads in the Americas-Caribbean and South America, 2013–2014. *Morbidity and Mortality Weekly Report* **63**, 500–501.
17. Thaikruea L *et al.* (1997) Chikungunya in Thailand: a re-emerging disease? *Southeast Asian Journal of Tropical Medicine and Public Health* **28**, 359–364.
18. Joint Investigation Team of Bureau of Epidemiology *et al.* (2008) Preliminary Report on Investigation of Chikungunya Fever Outbreak in Laharn Subdistrict, Yi-ngo district, Narathiwat Province, August–October 2008. Weekly Epidemiological Surveillance Report. pp. 717–722.
19. Powers AM *et al.* (2000) Re-emergence of chikungunya and O'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *Journal of General Virology* **81**(Pt 2), 471–479.
20. Rianthavorn P *et al.* (2010) An outbreak of chikungunya in southern Thailand from 2008 to 2009 caused by African strains with A226V mutation. *International Journal of Infectious Diseases* **14**(Supplement 3), e161–e165.
21. Theamboonlers A *et al.* (2009) Clinical and molecular characterization of chikungunya virus in South Thailand. *Japanese Journal of Infectious Diseases* **62**, 303–305.
22. Tssetsarkin KA *et al.* (2007) A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLOS Pathogens* **3**, e201.
23. Roiz D *et al.* Autochthonous chikungunya transmission and extreme climate events in Southern France. *PLOS Neglected Tropical Diseases* **2015**; **9**: e0003854.
24. Westbrook CJ *et al.* (2010) Larval environmental temperature and the susceptibility of *Aedes albopictus* Skuse (Diptera: Culicidae) to chikungunya virus. *Vector-Borne and Zoonotic Diseases* **10**, 241–247.
25. Delatte H *et al.* (2009) Influence of temperature on immature development, survival, longevity, fecundity, and gonotrophic cycles of *Aedes albopictus*, vector of chikungunya and dengue in the Indian ocean. *Journal of medical entomology* **46**, 33–41.
26. Gould EA and Higgs S (2009) Impact of climate change and other factors on emerging arbovirus diseases. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **103**, 109–121.

27. **Kumar NP *et al.*** (2008) A226V mutation in virus during the 2007 chikungunya outbreak in Kerala, India. *Journal of General Virology* **89**, 1945–1948.
28. **Thavara U *et al.*** (2009) Outbreak of chikungunya fever in Thailand and virus detection in field population of vector mosquitoes, *Aedes aegypti* (L.) and *Aedes albopictus* Skuse (Diptera: Culicidae). *Southeast Asian Journal of Tropical Medicine and Public Health* **40**, 951.
29. **Smith DL *et al.*** (2005) Assessing the role of long-distance translocation and spatial heterogeneity in the raccoon rabies epidemic in Connecticut. *Preventive Veterinary Medicine* **71**, 225–240.
30. **Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health** Available at <http://www.boe.moph.go.th/boedb/chikun/>. Accessed August 2010.
31. **Department of Provincial Administration, Ministry of Interior** (2010) Thailand Available at <http://stat.bora.dopa.go.th/stat/>. Accessed August 2010.
32. **Souris M** Available at <http://www.rsgis.ait.ac.th/~souris/thailand.htm#THAILANDUSE>. Accessed August 2010.
33. **Department of Agricultural Extension, Ministry of Agriculture and Cooperatives** (2010) Bangkok, Thailand. Available at <https://www.doae.go.th/doae/>. Accessed August 2010.
34. **Agricultural Research Development Agency** (2010) Bangkok, Thailand. Available at <http://www.arda.or.th/kasetinfo/south/para/controller/01-09.php>. Accessed August 2010.
35. **Goddard Earth Sciences Data and Information Services Center** (2016) Greenbelt, MD, USA. Available at https://disc.gsfc.nasa.gov/datacollection/TRMM_3B42RT_Daily_7.html. Accessed 2018.
36. **Huffman GJ and Bolvin DT.** *Real-time TRMM Multi-Satellite Precipitation Analysis Data Set Documentation*. NASA Technical Documents 2015.
37. **Huffman GJ *et al.*** (2007) The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology* **8**, 38–55.
38. **NASA EOSDIS Land Processes DAAC UEROaSEC, Sioux Falls, South Dakota**, Available at https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11a2. Accessed January 2018.
39. **Smith DL *et al.*** (2002) Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proceedings of the National Academy of Sciences* **99**, 3668.
40. **Kramer AM *et al.*** (2016) Spatial spread of the West Africa Ebola epidemic. *Royal Society Open Science* **3**, 160294.
41. **Google Inc** Available at <http://maps.google.com/>. Accessed 2010.
42. **Eggo RM, Cauchemez S and Ferguson NM** (2011) Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. *Journal of The Royal Society Interface* **8**, 233–243.
43. **Viboud C *et al.*** (2006) Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451.
44. **Nelder JA and Mead R** (1965) A simplex method for function minimization. *The Computer Journal* **7**, 308–313.
45. **Geem ZW** (2006) Parameter estimation for the nonlinear Muskingum model using the BFGS technique. *Journal of Irrigation and Drainage Engineering* **132**, 474–478.
46. **Safizadeh MH and Signorile R** (1994) Optimization of simulation via Quasi-Newton methods. *ORSA Journal on Computing* **6**, 398–408.
47. **Belik V, Geisel T and Brockmann D** (2011) Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review X* **1**, 011001.
48. **Robert MA *et al.*** (2016) Modeling mosquito-borne disease spread in U.S. Urbanized areas: the case of dengue in Miami. *PLOS ONE* **11**, e0161365.
49. **Tran A *et al.*** (2013) A rainfall- and temperature-driven abundance model for *Aedes albopictus* populations. *International Journal of Environmental Research and Public Health* **10**, 1698.
50. **Paily KP *et al.*** (2013) Efficacy of a mermithid nematode *Romanomermis iyengari* (Welch) (Nematoda: Mermithidae) in controlling tree hole-breeding mosquito *Aedes albopictus* (Skuse) (Diptera: Culicidae) in a rubber plantation area of Kerala, India. *Parasitology Research* **112**, 1299–1304.
51. **Sumodan PK *et al.*** (2015) Rubber plantations as a mosquito box amplification in South and Southeast Asia. In Morand S *et al.* (eds), *Socio-Ecological Dimensions of Infectious Diseases in Southeast Asia*. Singapore: Springer Singapore, pp. 155–167.