Inter-annual variation in seasonal dengue epidemics driven by multiple interacting factors in Guangzhou, China

Rachel J. Oidtman[1], Shengjie Lai[2,3,4], Zhoujie Huang[2], Amir S. Siraj[1], Robert C. Reiner[5], Andrew J. Tatem[3,4], T. Alex Perkins*[1], Hongjie Yu*[2]

[1] Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA
[2] School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China
[3] WorldPop, Department of Geography and Environment, University of Southampton, Southampton, UK
[4] Flowminder Foundation, Stockholm, Sweden
[5] Institute for Health and Metrics and Evaluation, University of Washington, Seattle, WA, USA

* Corresponding authors: yhj@fudan.edu.cn and taperkins@nd.edu

ABSTRACT

Vector-borne diseases display wide inter-annual variation in seasonal epidemic size due to their complex dependence on temporally variable environmental conditions and other factors. In 2014, Guangzhou, China experienced its worst dengue epidemic on record, with incidence exceeding the historical average by two orders of magnitude. To disentangle contributions from multiple factors to inter-annual variation in epidemic size, we fitted a semi-mechanistic model to time series data from 2005-2015 and performed a series of factorial simulation experiments in which seasonal epidemics were simulated under all combinations of year-specific patterns of four time-varying factors: imported cases, mosquito density, temperature, and residual variation in local conditions not explicitly represented in the model. Our results indicate that while epidemics in most years were limited by unfavorable conditions with respect to one or more factors, the epidemic in 2014 was made possible by the combination of favorable conditions for all factors considered in our analysis.

In response to warming temperatures and other consequences of climate change, many regions are becoming increasingly suitable for pathogens whose transmission is sensitive to climatic conditions[1,2]. One clear example is El Niño, which is known to drive inter-annual variation in epidemics of both water-borne[3,4] and vector-borne[5,6] pathogens. Another example is rising temperatures, which have been hypothesized to facilitate pathogen range expansion[7] and to fuel more explosive epidemics[8].

Although there are clear links between local climatic conditions and the transmission of numerous pathogens, those links can be difficult to isolate in epidemiological analyses[9,10]. In an area where a pathogen is endemic, its host population may often be considerably immune, resulting in inter-annual variation driven by a combination of time-varying climatic conditions and nonlinear feedbacks of population immunity[11–13]. Those feedbacks may furthermore vary with shifting patterns of host demography, which modulate ebbs and flows in the pool of susceptible hosts over time[14]. In an area where a pathogen is not endemic but is instead transmitted in the context of limited seasonal epidemics, pathogen importation can play a critical role in limiting, or enabling, transmission[15].

The recent history of dengue virus (DENV) in the city of Guangzhou, China presents an ideal opportunity to examine how temporal variation in local climatic conditions and pathogen importation interact to drive inter-annual variability in transmission in a seasonally epidemic context. Since 1990, mainland China has experienced highly variable, but relatively low DENV transmission, with a median of 376 and a range of 2-6,836 cases reported annually from 1990-2004[16]. More recently, increasingly large seasonal epidemics have occurred, with a median of 438 and a range of 59-47,056 cases from 2005-2014[16]. These epidemics have been highly seasonal and distinct from year to year, given the markedly seasonal climatic conditions in portions of mainland China where dengue is locally transmitted. At the same time, the endemic status of DENV in neighboring southeast Asia ensures a reliable, and growing, source of DENV importation into Guangzhou and elsewhere in mainland China[17].

Following the epidemic of 37,445 locally acquired dengue cases in Guangzhou in 2014, there has been growing interest in modeling DENV transmission there. Different models have emphasized different drivers, however, leading to different conclusions. Two studies concluded that climatic conditions were the primary driver of DENV transmission[18,19], whereas others concluded that importation patterns, combined with delayed outbreak response, were causal drivers of the 2014 outbreak[20–22]. Still others found that neither climatic conditions nor importation were the key drivers of transmission, but instead that urbanization was pivotal[23,24]. Several analyses that used incidence data aggregated at a monthly time scale for 2005-2015 showed high predictive capability at one-month lead times[18,19] but did not facilitate clear interpretation of how importation interacts with local conditions to result in high inter-annual variation in transmission. Mechanistic models applied to date have used daily or weekly data, but only for 2013-2014, and therefore only considered years with anomalously high transmission[20,22–24]. As a result, it is unclear how well those models could explain the strikingly low incidence observed in years other than 2013-2014.

Here, we applied a stochastic, time-series susceptible-infected-recovered (TSIR) model[25] that we fitted to daily dengue incidence data from 2005-2015 to determine the relative roles of local conditions and pathogen importation in driving inter-annual variation in DENV transmission. To make detailed inferences with incidence data at daily resolution, we made several enhancements to the standard TSIR framework, including a realistic description of the DENV generation interval, lagged effects of covariates on transmission, and flexible spline relationships between covariates and their contributions to transmission. After fitting the model

and checking its consistency with the data to which it was fitted, we conducted simulation experiments in which we examined how annual incidence of locally acquired dengue differed across simulations with inputs about local conditions and importation patterns that varied from year to year. Using a series of factorial simulation experiments, we quantified the relative contributions of time-varying local factors and importation to inter-annual variation in dengue incidence.

RESULTS

Estimating relationships between local conditions and transmission

In the 11-year time series that we examined, there was marked seasonal variation in local dengue incidence and in putative drivers of DENV transmission (Fig 1). We estimated a latent mosquito density curve, $m(t)$, using two different types of entomological data (Fig 1 C), resulting in a seasonal pattern of mosquito density. Although inter-annual variation in mosquito density and temperature was minimal (Fig 1 C, D), there was pronounced inter-annual variation in imported and local dengue incidence (Fig 1 A, B).

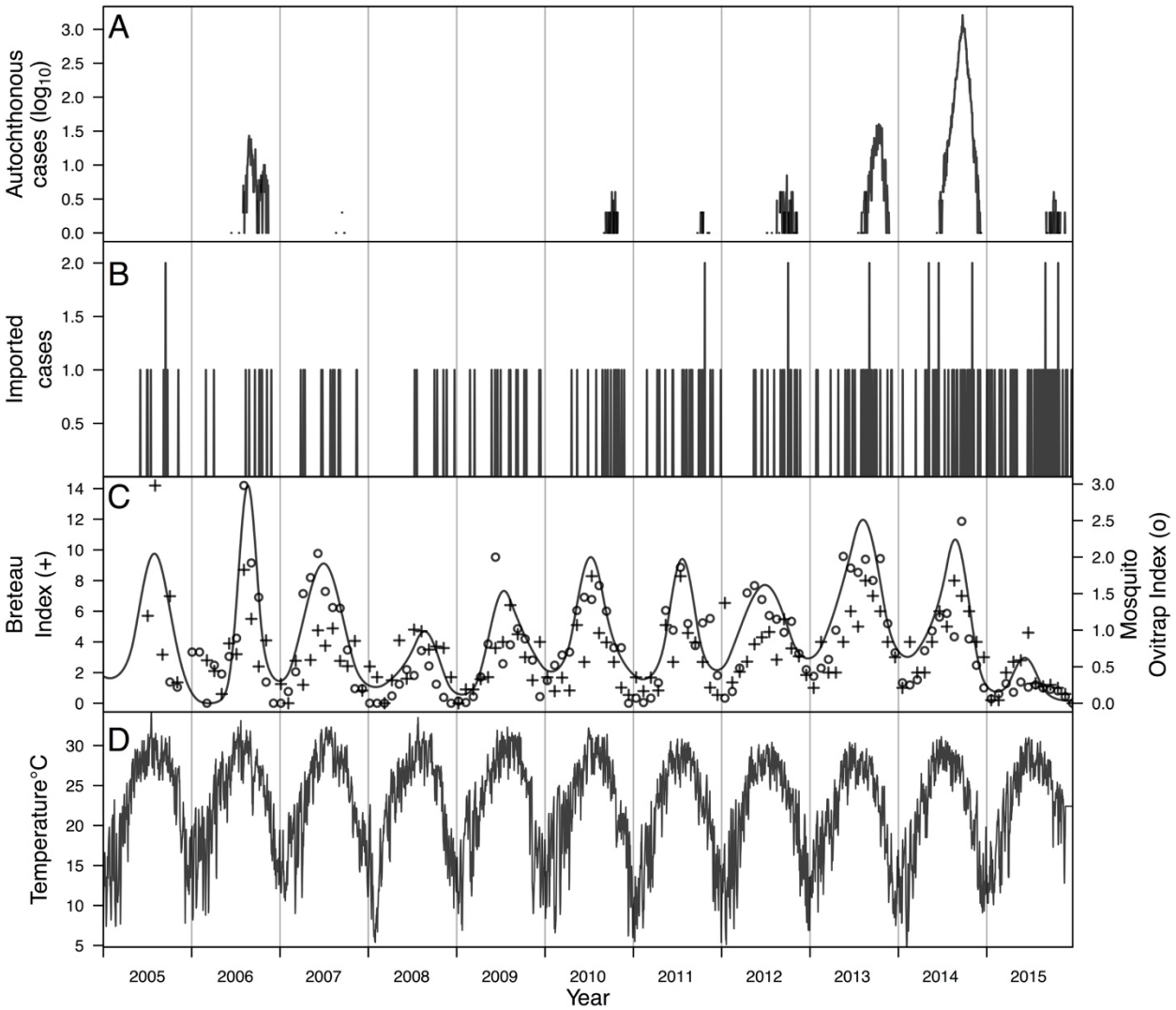

Fig 1. Time series of data from 2005-2015 in Guangzhou, China. (A) Local dengue incidence. (B) Imported dengue incidence. (C) Breteau index (BI) of mosquito density, mosquito

ovitrap index (MOI) of mosquito larval density, and maximum likelihood estimate of latent mosquito density variable, *m*(*t*). (D) Daily mean temperature.

To understand the association between temporal variation in local conditions and local DENV transmission, we fitted two bivariate basis functions with cubic B-splines that allowed for distinct lagged relationships between temperature, mosquito density, and their effects on transmission (Fig 2). For temperature, we found that its contribution on a given day to the time-varying transmission coefficient, $\beta(t)$, generally peaked near 30 °C and that the magnitude of these daily contributions was somewhat larger at shorter lags (Fig 2A). For mosquito density, we found a positive relationship between $m(t)$ and $\beta(t)$ at all lags, with the daily contribution of mosquito density at intermediate lags being 10-15% larger than at shorter or longer lags (Fig 2B). In addition, we estimated a term, $\beta_0(t)$, that explicitly modeled residual variation in $\beta(t)$ that was not accounted for by temperature or mosquito density but that was necessary to reproduce observed patterns of local dengue incidence (Fig 3). The 95% posterior predictive interval for this term was entirely positive during the transmission season in 2014, whereas in other years it fluctuated relatively tightly around zero (Fig 3B). This implies that appealing to systematic differences with respect to one or more unspecified local conditions is necessary to explain the anomalously high incidence observed in 2014.

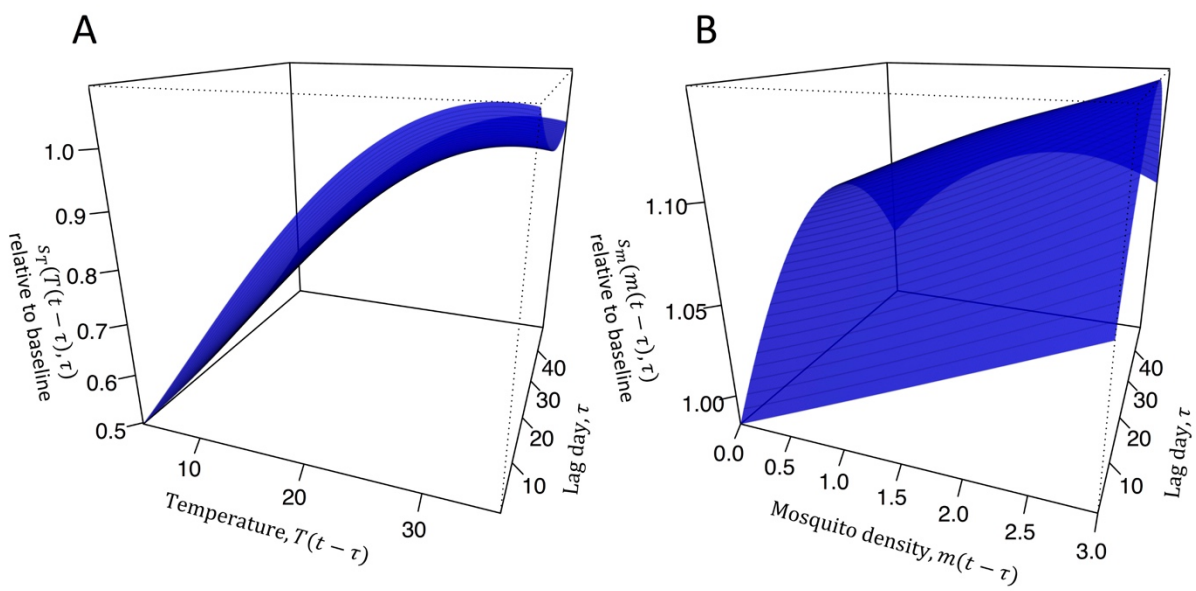

Fig 2. Contributions of temperature (A) and mosquito density (B) at different lags τ to the transmission coefficient on day *t*. Following eqn. (4), the surface in A corresponds to $s_T(T_{t-\tau},\tau)$ and the surface in B corresponds to $s_m(m(t-\tau),\tau)$, both of which are summed across values of τ ranging 1-49, exponentiated, and multiplied by each other and $e^{\beta_0(t)}$ to obtain $\beta(t)$. Values of parameters informing these surfaces shown here represent medians from the posterior distribution of parameters.

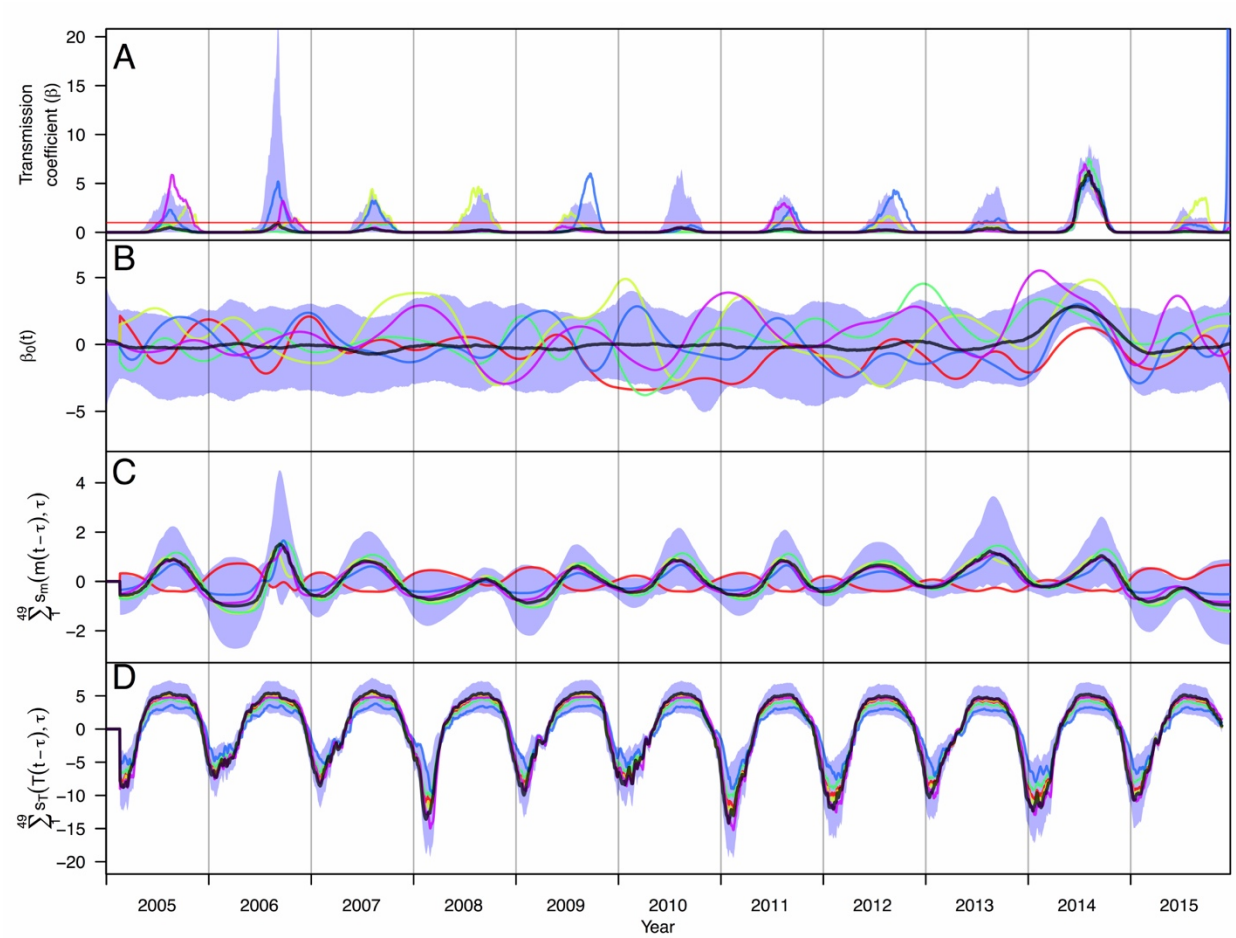

Fig 3. Time series of posterior estimates of the time-varying transmission coefficient $\boldsymbol{\beta(t)}$ (A) and contributions thereto from residual local conditions, $\boldsymbol{\beta_0(t)}$ (B), mosquito density, $\boldsymbol{m(t)}$ (C), and temperature (D). Different colored lines correspond to different samples from the posterior distribution of parameter values, which provide information about correlations among parameters that pertain to different components of the model. The red horizontal line in A indicates $\beta(t) = 1$. The shaded blue region represents the 95% posterior predictive interval, and the black line is the median value.

By estimating three separate components of $\beta(t)$, we were able to evaluate the relative contributions to $\beta(t)$ of each of mosquito density, temperature, and other local conditions at different points in time (Fig 3). We found that mosquito density tended to have a smaller but more variable effect compared to temperature, which resulted in considerably lowered $\beta(t)$ values at low temperatures. In most years, the effect of mosquito density on $\beta(t)$ tended to be more pronounced within a shorter seasonal time window than did that of temperature (Fig 3C, D). The contributions of $\beta_0(t)$ to $\beta(t)$ were highly variable across different draws from the posterior, other than the consistently large, positive effect in 2014 (Fig 3B).

Checking model consistency with data

We used data on imported cases to seed 1,000 simulations of local DENV transmission over the entire 2005-2015 time period, with local transmission patterns in each simulation determined by a different random draw from the posterior distribution of model parameters. Over the period as a whole, daily medians of simulated local dengue incidence were highly correlated ($\rho = 0.966$) with observed local incidence (Fig 4A). Within each year, observed "features" of local dengue

incidence patterns were generally consistent with simulations. For annual incidence across years and peak weekly incidence, observed values fell within the 95% posterior predictive intervals (PPI) in 11/11 years (Fig S1, S2). For total number of weeks with non-zero incidence and for the length of the transmission season (time elapsed between first and last cases), observed values fell within the 95% PPIs in 8/11 and 9/11 years, respectively (Fig S3, S4). Years for which observed values fell outside of the 95% PPI tended to be those with intermediate levels of transmission (2006, 2013) or longer transmission seasons (2010, 2015). Years for which observed values most consistently fell within the 95% PPIs were those with either low (2007, 2008) or high (2014) transmission. In addition, we found that the fitted model correctly ranked 2014 as the year with the highest annual incidence 100% of the time (Fig 5). Other notable high years included 2013 and 2006, which were both correctly ranked as years with relatively high local incidence. Years with low local incidence that our model ranked correctly included 2008-2011 and 2015.

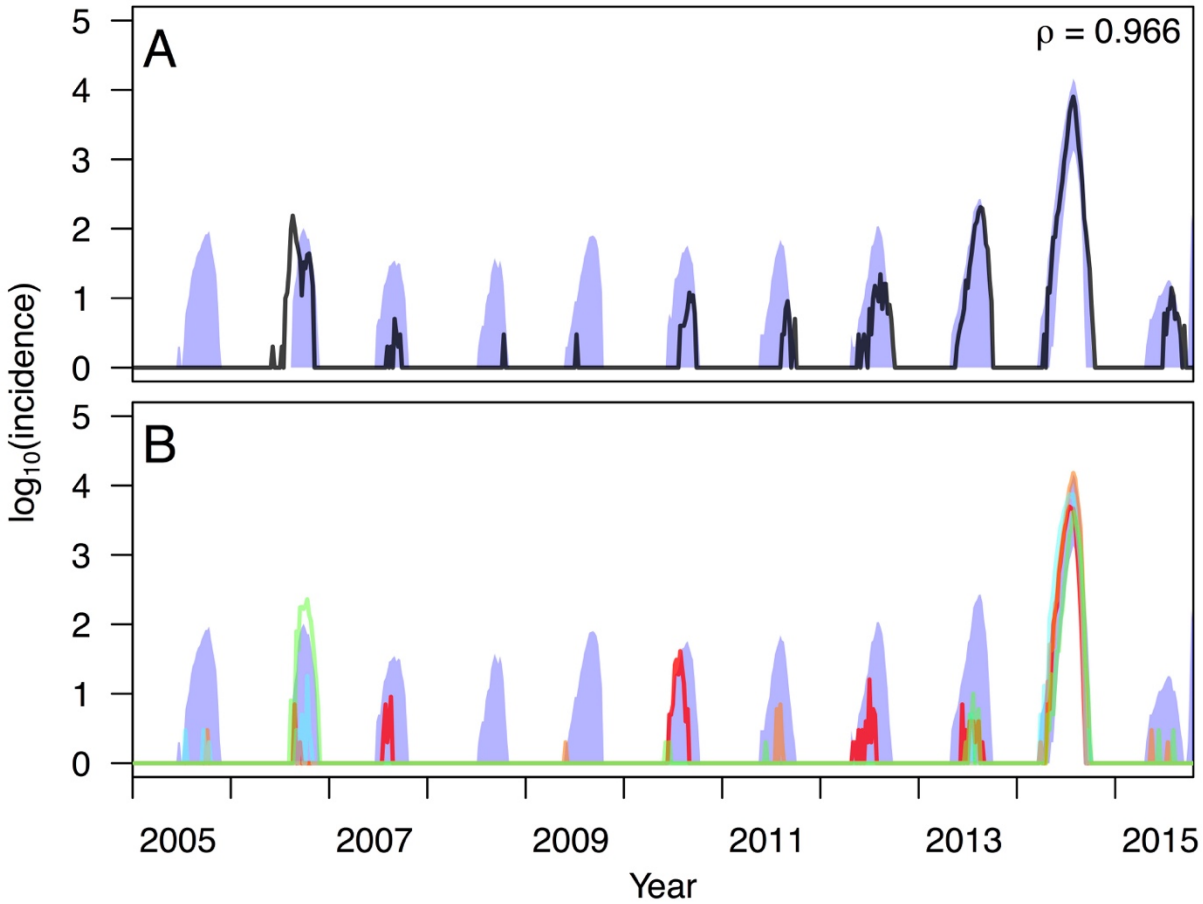

Fig 4. Correspondence between empirical and simulated patterns of local dengue incidence. In A, the black line shows empirical values of log local incidence and the blue band shows the 95% posterior predictive interval from model simulations. The value of Pearson's correlation coefficient indicated in the upper right pertains to untransformed daily values between model simulations and data. In B, different colored lines correspond to simulations based on different samples from the posterior distribution of parameter values. Simulations of local transmission were seeded by data on imported cases and otherwise used the fitted model of local transmission to simulate local cases.

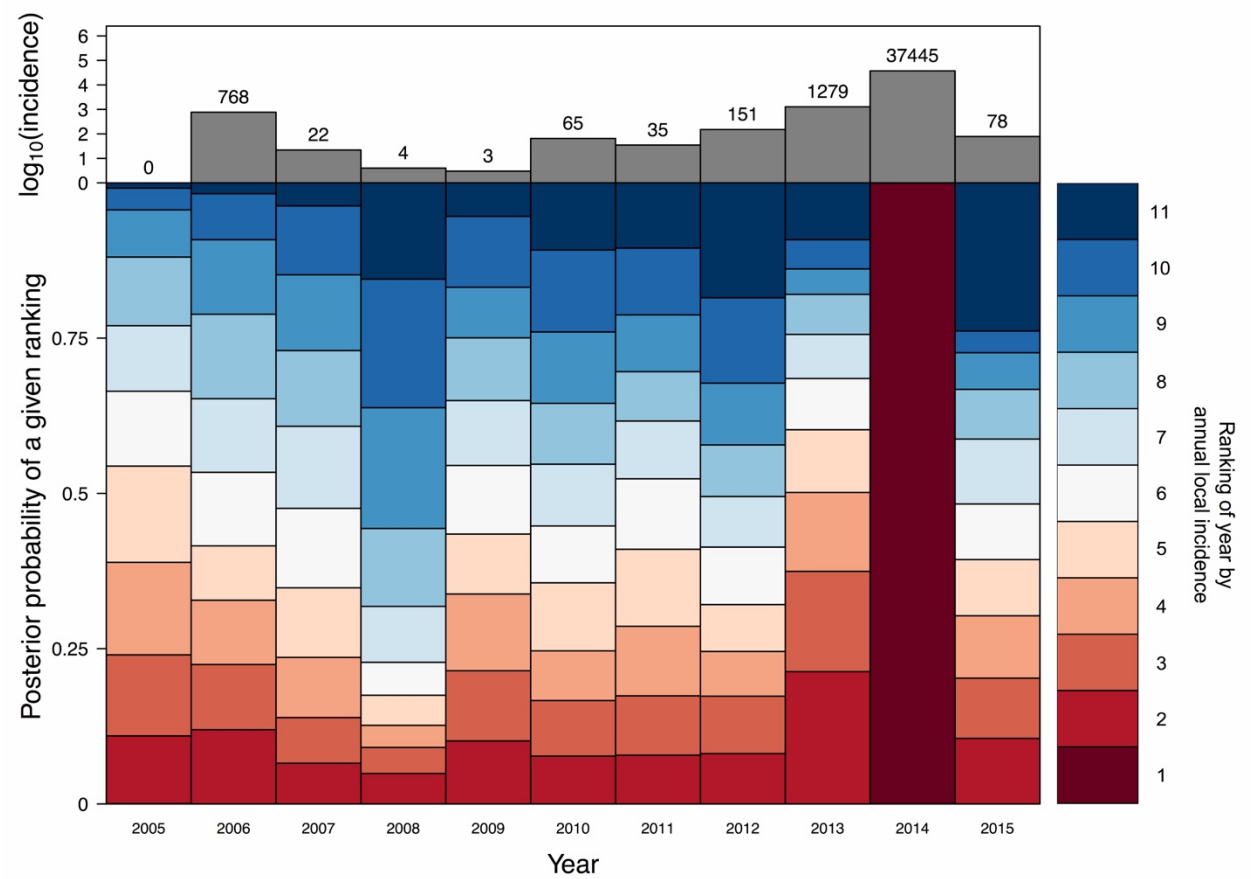

Fig 5. Ranking of years by annual local incidence simulated with the fitted model given data on local conditions and imported cases from that year. Dark red corresponds to the lowest ranking (i.e., highest simulated local incidence), and dark blue corresponds to the highest ranking (i.e., lowest simulated local incidence). The height of a given segment of a given year's bar is proportional to the posterior probability that simulations of local incidence from that year were of a given rank relative to other years. For example, the dark red segments in 2013 and 2014 indicate a high posterior probability that model simulations correctly resulted in high rankings for those years. The total number of dengue cases reported in each year is shown in the top panel for reference.

Disentangling drivers of inter-annual variation in dengue incidence

Once we determined that simulations from the fitted model were consistent with observed patterns, we performed a factorial simulation experiment in which we swapped local conditions from each year with imported case patterns from each other year and used those conditions to drive simulations of 1,000 replicate transmission seasons under each of these 122 combinations. Some of the more extreme contrasts illustrate the reasoning behind this approach. For example, given local conditions in 2014, our model projects that much higher local incidence would have been observed under importation conditions experienced in most other years (Fig 6, top). In contrast, given imported case patterns from 2008, our model projects that very low local incidence would have resulted from local conditions in every year, including 2014 (Fig 6, bottom).

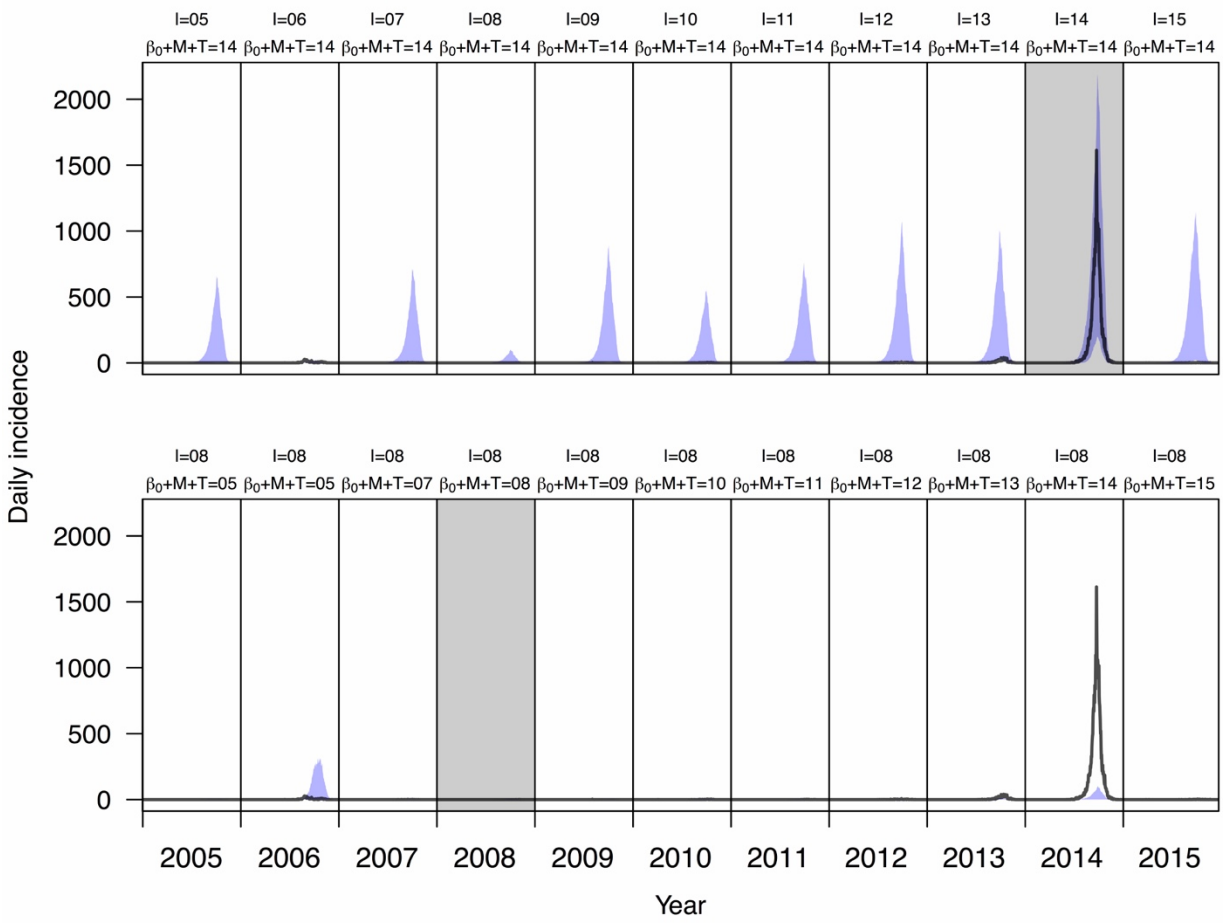

Fig 6. Example outputs from the factorial simulation experiments. The top row shows a series of 1,000 simulations per year (blue = 95% posterior predictive interval) combining local conditions from 2014 with importation patterns from each year in 2005-2015. The bottom row shows a series of 1,000 simulations per year combining importation patterns from 2008 with local conditions from each year in 2005-2015. These examples contrast seasonal epidemic patterns that would have resulted from hypothetical situations swapping year-specific local conditions and year-specific importation patterns from different years. For reference, empirical patterns of local incidence are shown with black lines.

Of variation in simulated local incidence that was accounted for by the model (51.2%) in a two-way analysis of variance, local conditions accounted for approximately 88.9% and imported case patterns accounted for the remaining 11.1% (Table 1, first row). A large amount of residual variation (49.8%) in simulated local incidence across replicates was not explained by either factor and instead reflected the highly stochastic nature of epidemics in this setting where DENV transmission is so volatile. Because the number of replicates was at our discretion in this simulation experiment, the p-value from this analysis of variance was not meaningful[26]. On the whole, these results showed that high local incidence was unlikely to occur in years in which local conditions were not highly suitable for transmission (Fig 7). Performing a similar factorial simulation experiment in which we dropped 2014, we found that local conditions only accounted for 38.3% of variation in local incidence explained by the model, while importation accounted for 61.7% (Table 1, second row).

Years included	SSQ_{β_0+M+T} (% Model)	SSQ_I (% Model)	SSQ_{res} (% Total)	SSQ Total
2005-2015	4.1×10^5 (88.9)	5.1×10^4 (11.1)	4.6×10^5 (49.8)	9.1×10^5
2005-2013, 2015	1.7×10^4 (38.3)	2.8×10^4 (61.7)	3.7×10^5 (88.9)	4.1×10^5

Table 1. Sum of squared error (SSQ) and normalized percent variation described by local conditions ($\beta_0 + M + T$), imported case patterns (I), or residual stochasticity in the two-way analysis of variance on the factorial simulation experiments. The % Model values were calculated by dividing the SSQ explained by a given variable by the sum of SSQ values from all four model variables.

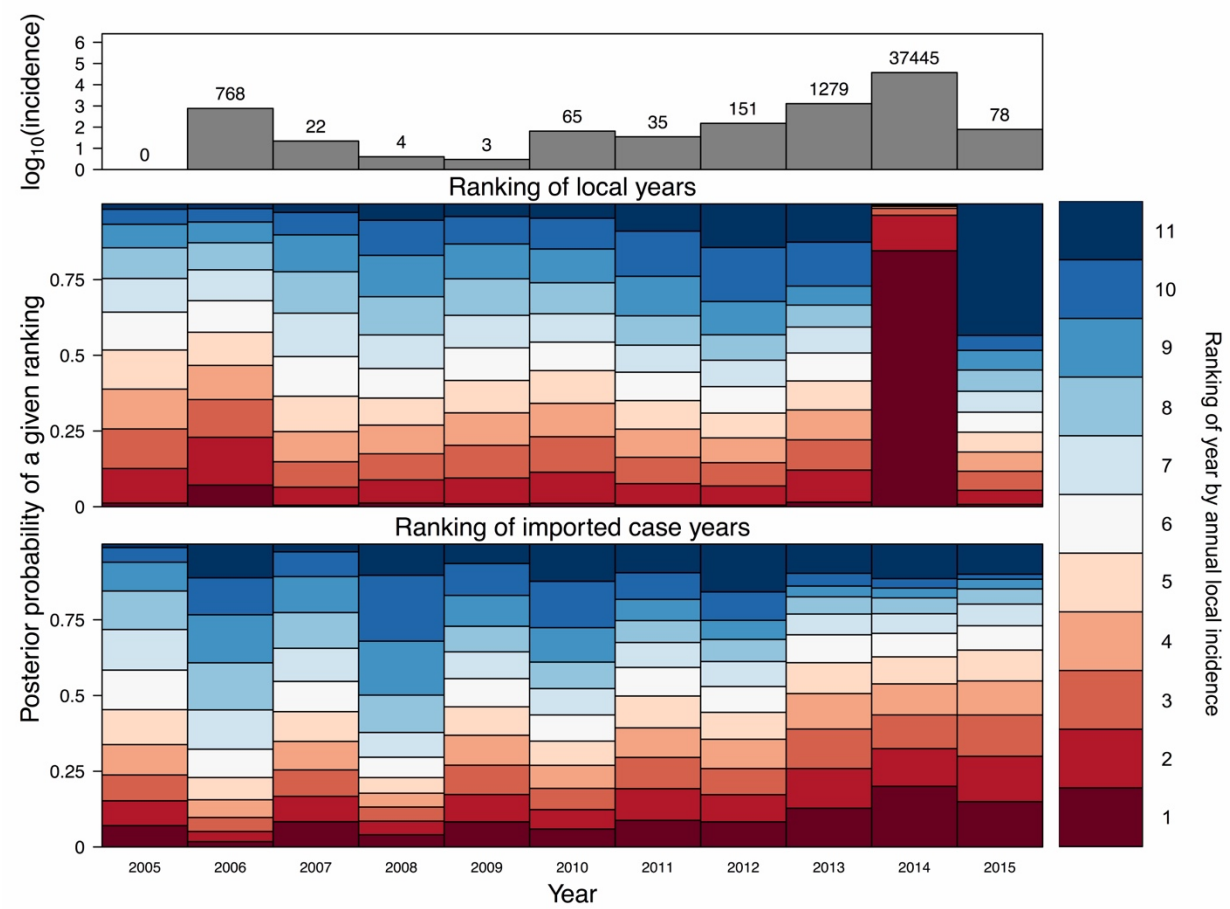

Fig 7. Ranking of years by simulated annual incidence from the factorial simulation experiment. (Top) The total number of dengue cases reported in each year is shown in the top panel for reference. (Middle) Simulations with the fitted model given data on local conditions from a given year and imported cases from each of the other years. (Bottom) Simulations with the fitted model given data on imported cases from a given year and local conditions from each of the other years. A large amount of red in a column indicates that conditions in that year were relatively favorable for transmission, whereas blue indicates the opposite, similar to Fig. 5.

To parse the individual contributions of each local variable to inter-annual variation in local incidence, we performed an additional factorial simulation experiment in which we swapped all possible combinations of imported case patterns, mosquito density, temperature, and β_0 from different years. We performed a set of 1,000 replicate simulations for the 14,641 possible ways that year-specific patterns could be combined, allowing us to account for possible interactions

among these four variables. Calculating the variation explained by β_0 in the four-way analysis of variance, we found that it accounted for 75.4% of all variation in local incidence accounted for by the model (Table 2). Imported cases contributed the next largest portion of variation (11.3%), whereas mosquito density (9.5%) and temperature (3.8%) each contributed less. Repeating this four-way analysis of variance without data from 2014, the proportions of all variation explained by the models were somewhat more consistent across these four variables (Table 2, second row). Furthermore, the rankings of relative contributions of the four variables changed, with mosquito density (35.9%) and imported cases (39.2%) playing dominant roles in driving inter-annual variation in local incidence.

Years included	SSQ_{β_0} (% Model)	SSQ_M (% Model)	SSQ_I (% Model)	SSQ_T (% Model)	$\mathrm{SSQ}_{\mathrm{res}}$ (% Total)	SSQ Total
2005-2015	4.2×10^7 (75.4)	5.2×10^6 (9.5)	6.3×10^6 (11.3)	2.1×10^6 (3.8)	7.8×10^7 (58.7)	1.3×10^8
2005-2013, 2015	7.6×10^5 (10.5)	2.6×10^6 (35.9)	2.9×10^6 (39.2)	1.0×10^6 (14.3)	5.0×10^7 (87.2)	5.7×10^7

Table 2. Sum of squared error (SSQ) and normalized percent variation described by the predictor variables and the residuals in the four-way analysis of variance on the factorial simulation experiments. The % Model values were calculated by dividing the SSQ explained by a given variable by the sum of SSQ values from all four model variables.

DISCUSSION

Populations subject to seasonal epidemics of any number of diseases are prone to high variability in epidemic size, due to inter-annual variation in imported cases that seed those epidemics and inter-annual variation in local conditions that drive transmission. We estimated the relative contributions of local conditions and importation in driving inter-annual variation in dengue epidemics in Guangzhou, China, which has recently been subject to seasonal epidemics ranging four orders of magnitude in size. Other studies[18–23] have investigated the same 11-year time series, either in whole or in part, but arrived at differing conclusions and did not take full advantage of the exceptional level of detail in this data set (Table S1). By leveraging these data more fully and using a modeling framework that blends elements of mechanistic and statistical modeling, we showed that local conditions and importation patterns jointly determined epidemic size in most years and that anomalous local conditions were responsible for one anomalously large epidemic. Specific examples from this 11-year time series reinforce the notion that either or both of these factors can limit epidemic size (Fig 7).

Regarding the large epidemic in 2014, our results suggest that unknown local conditions captured by β_0 played a dominant role in driving this extreme event. For this reason, the unknown factors that β_0 was picking up on should now be of great interest. One possibility is that transmission was actually not much higher in 2014 but instead that a larger proportion of DENV infections resulted in symptomatic disease in 2014 than in other years. This could have occurred if a large number of people experienced a mild or asymptomatic first infection prior to 2014 and then experienced a more severe second infection in 2014[27]. It is also possible that the DENV serotype or genotype that circulated in 2014 could have been more infectious[28] or more virulent[29,30]. Another possibility is that, although mosquito densities were not notably higher in 2014 than in other years, there could have been undetected changes in the composition of the mosquito population that enhanced their competence for transmission. Demographic dynamics are known to result in substantial temporal variation in the age profile of mosquito populations[31], which has been shown to be a key determinant of dengue epidemic size in some settings[32]. Yet one more possibility is that media attention during the 2014 epidemic[33,34] could have heightened

awareness of dengue and led to an increase in the number of people recorded by the surveillance system[35]. Without further investigation, these hypotheses all remain plausible, with each potentially acting alone or in combination to contribute to the unprecedented dengue incidence observed in 2014.

Although it is somewhat unsatisfying that our analysis could not pinpoint the cause of the 2014 epidemic more specifically, clearly defining the roles that known factors played and ruling them out as primary drivers of the 2014 epidemic is also of great value. Our results did show that temperature played a role in delimiting the transmission season, that mosquito density influenced the timing and extent of transmission within a season, and that importation regulated the potential for local transmission in a given season. Our modeling approach was unique in allowing us to isolate each of those effects by building on prior knowledge about them in such a way that we captured their differential influence at different lags and captured the extent to which imported dengue cases translated into locally acquired cases. Had we fitted our model solely to data from 2013 and 2014, as others have[20,22–24], we likely would have misestimated the contributions of these factors to local transmission and would not have been able to detect the anomalous local conditions in 2014 that appear to have driven the large epidemic that year.

Our model incorporates a number of innovations that were essential for obtaining our results, including the ability to accommodate daily incidence data, to adapt the timescale of transmission to the pathogen's generation interval, to estimate multiple lagged effects in a flexible manner, and to isolate the timing of residual variation in transmission, all of which may prove useful to time series analyses of climate-sensitive pathogens[10,36]. At the same time, there are important limitations of our approach. First, even though it is well known that many DENV infections are inapparent[27], we worked under the assumption that cases detected through passive surveillance were representative of the true incidence of infection. Combining data augmentation methods[37] with hypotheses about ways in which reporting rates might vary through time could offer one way to relax this assumption. Second, we assumed that the population was immunologically naïve and remained so over time. The limited data available pertaining to this question suggest that DENV immunity is indeed low (2.43%, range: 0.28-5.42%)[38], meaning that impacts of immunity on transmission should be negligible. These effects could be stronger at finer spatial scales, however.

In conclusion, our finding that epidemic size in any given year depends on a complex interaction between importation and local conditions suggests that public health authorities should not focus on only one of these factors at the exclusion of others. As some studies have done[20–22], it is tempting to attribute the increase in local dengue incidence in Guangzhou to the concurrent increase in imported dengue. Our results suggest that doing so belies the important role that local conditions play in limiting or enhancing transmission in any given year. What an overly simplistic view risks is allowing for another epidemic like the one in 2006, which our results suggest was driven by favorable local conditions despite relatively low importation. Moreover, understanding and reducing the favorability of local conditions for transmission may also mean the difference between years like 2014 and 2015, with importation high in both years but local transmission much lower in 2015. Given the global expansion of DENV and other viruses transmitted by *Aedes* mosquitoes, improved understanding of the interactions among multiple drivers in settings with potential for seasonal DENV transmission—including portions of Australia, the United States, and the Mediterranean—will be essential for reducing the risk of large epidemics such as the one observed in Guangzhou in 2014.

METHODS

Data

Data on locally acquired and imported dengue cases from 2005-2015 were obtained from the Health Department of Guangdong Province (http://www.gdwst.gov.cn/). As one of the statutorily notifiable infectious diseases in China, dengue is diagnosed according to national surveillance protocol with standardized case definitions described in detail elsewhere[16]. An imported case is defined as one for which the patient had traveled abroad to a dengue-endemic country within 15 days of the onset of illness. In some cases, importation is defined based on laboratory results showing that the infecting DENV had a high sequence similarity in the preM/E region compared with viruses isolated from the putative source region where the patient had traveled. Otherwise, a dengue case was considered locally acquired. This determination was made by local public health institutes. All the data used in this study were anonymized; the identity of any individual case cannot be uncovered.

We utilized indices of both adult mosquito density and larval density, which are available from the Guangzhou Center for Disease Control and Prevention (http://www.gzcdc.org.cn). Adult *Aedes albopictus* mosquitoes were sampled by light traps with mosquito ovitrap index (MOI). The MOI was defined as the number of positive ovitraps for adult and larval *Ae. albopictus* per 100 retrieved traps[39]. Breteau index (BI), which measures the density of *Ae. albopictus* mosquito larvae, was the number of positive containers per 100 houses inspected[39,40]. Both indices were measured monthly and comprised the information on mosquito density sampled in residential households (>50 households sampled per month), parks, construction sites, and other urban areas. Data on daily average temperature were obtained from the China Meteorological Data Sharing Service System (http://data.cma.cn).

Model description

Transmission modeling framework

A general framework that can be used to model the relationship between cases from one generation to the next is the TSIR model[24]. Under an adaptation of that model to realistically account for time lags associated with vector-borne pathogen transmission[41], I_t is defined as the number of new local cases at time *t* and I'_t is the effective number of cases, both local and imported, that could have generated a local case at time *t*. This effective number of cases in the previous generation is defined as

$$I'_t = \sum_{\tau=1}^{49} \omega_\tau (I_{t-\tau} + \iota_{t-\tau}), \quad (1)$$

where ω_τ is the probability that the serial interval is τ days[41]. We described ω_τ with a function derived by Siraj et al.[8] based on first-principles assumptions about DENV transmission. The flexibility afforded by eqn. (1) allowed us to fully utilize the daily resolution of case data available for Guangzhou, which distinguished between imported and local cases, ι_t and I_t, respectively.

Consistent with other TSIR models, the relationship between I'_t and I_t was assumed to take the form

$$I_t = \beta(t) \frac{I'_t}{N} S'_t, \quad (2)$$

where $\beta(t)$ is the transmission coefficient on day t, N is population size, and S'_t is the number of susceptible people who could potentially become infected and present on day t. Due to the low incidence of dengue in Guangzhou on a per population basis (40,108 cases detected by surveillance during 2005-2015 in a city of 14 million), the number of susceptible people at any given time changes very little and remains very close to the overall population size. Therefore,

we assumed that $S'_t \approx N$, meaning that these terms canceled out in eqn. (2). Also because of such low incidence, including many days with zero incidence, accounting for the role of stochasticity in transmission was essential. Eqn. (2) has a clear and direct stochastic analogue in

$$I_t \sim \text{negative binomial}(\beta(t)I'_t, I'_t), \qquad (3)$$

where $\beta(t)I'_t$ is the mean parameter and I'_t is the clumping parameter of the negative binomial distribution[42,43].

Drivers of local transmission

We assumed that the potential for local transmission at time t, represented by $\beta(t)$, was determined by a combination of a latent variable representing mosquito abundance at time t, $m(t)$, temperature at time t, T_t, and other factors not accounted for directly by available data, such as mosquito control or preventative measures taken by local residents. Although the role of these factors in driving transmission is commonly assumed by models[44] and consistent with the highly seasonal nature of DENV transmission in Guangzhou[16], it is also clear that these factors may influence transmission considerably in advance of a case occurring. For example, high mosquito densities would be expected to affect transmission 2-3 weeks in advance, rather than instantaneously, to allow mosquitoes sufficient time to become infected, incubate the virus, and transmit it[8].

To afford the model sufficient flexibility to account for such lagged effects, we allowed $\beta(t)$ to depend on weighted sums of daily effects of $m(t-\tau)$ and $T_{t-\tau}$ for $\tau \in \{1, \ldots, 49\}$, which spanned the full range of serial intervals that we assumed were possible. Because the effects of $m(t-\tau)$ and $T_{t-\tau}$ could differ for different values of τ in complex ways, we defined flexible bivariate basis functions $s_m(m(t-\tau), \tau)$ and $s_T(T_{t-\tau}, \tau)$ with cubic B-splines to capture the contribution of daily conditions on day $t-\tau$ to $\beta(t)$ using the fda package in R[45]. Each of $s_m(m(t-\tau), \tau)$ and $s_T(T_{t-\tau}, \tau)$ was defined by nine parameters associated with a 3x3 matrix that defined the height of each component of the bivariate spline ranging 1-49 days for τ, 4-36 °C for T, and 0-5 for m, with the units of the latter corresponding to the scale of the mosquito oviposition index. That particular choice of units was not of consequence to the model, however, because a different choice would simply result in different values of parameters in $s_m(m(t-\tau), \tau)$ but yield the same values of $\beta(t)$. These lagged daily effects combined to define

$$\beta(t) = e^{\sum_{\tau=1}^{49} s_T(T_{t-\tau},\tau)} e^{\sum_{\tau=1}^{49} s_m(m(t-\tau),\tau)} e^{\beta_0(t)}, \qquad (4)$$

where $\beta_0(t)$ is a univariate cubic B-spline function that defines the time-varying contribution of factors other than temperature and mosquito abundance to $\beta(t)$. We specified $\beta_0(t)$ as a univariate spline with three evenly spaced knots per year across the 11-year time period, requiring a total of 33 parameters. We also represented the latent mosquito density variable $m(t)$ using a univariate cubic B-spline function with three knots per year for the 11-year time period. This variable allowed us to reconcile differences between the MOI and BI mosquito indices and to obtain daily values for mosquito abundance based on monthly indices.

Model fitting

We used a two-step process to estimate the posterior probability distribution of model parameters. First, we fitted the entomological model (i.e., $m(t)$) using maximum likelihood. Second, we fitted the epidemiological model using a Sequential Monte Carlo (SMC) algorithm in the BayesianTools R library[46].

Entomological likelihood

The probability of the full mosquito index time series, $\overrightarrow{\text{MOI}}$ and $\overrightarrow{\text{BI}}$, depends on the 33 parameters that define $m(t)$ (referred to collectively as $\vec{\theta}_m$) and three parameters, μ_{BI}, σ_{BI}, and

σ_{MOI}, that define an observation model relating $m(t)$ to the data. Under this model, the probabilities of these data are

$$\Pr\left(\overrightarrow{\mathrm{MOI}}|\vec{\theta}_m, \sigma_{MOI}\right) = \prod_t \phi(\mathrm{MOI}_t|\overline{m}(t), \sigma_{MOI}) \quad (5)$$

and

$$\Pr\left(\overrightarrow{\mathrm{BI}}|\vec{\theta}_m, \mu_{BI}, \sigma_{BI}\right) = \prod_t \phi(\mathrm{BI}_t|\mu_{BI}\overline{m}(t), \sigma_{BI}), \quad (6)$$

where $\phi(x|\mu, \sigma)$ denotes a normal probability density with parameters μ and σ evaluated at x and $\overline{m}(t)$ denotes the monthly average of $m(t)$. Together, eqns. (5) and (6) specify

$$\mathcal{L}\left(\vec{\theta}_m, \sigma_{MOI}, \mu_{BI}, \sigma_{BI}\middle|\overrightarrow{\mathrm{MOI}}, \overrightarrow{\mathrm{BI}}\right) = \Pr\left(\overrightarrow{\mathrm{MOI}}|\vec{\theta}_m, \sigma_{MOI}\right) \Pr\left(\overrightarrow{\mathrm{BI}}|\vec{\theta}_m, \mu_{BI}, \sigma_{BI}\right), \quad (7)$$

which is the overall likelihood of the entomological model parameters.

Entomological model fitting

We obtained maximum-likelihood estimates of $\vec{\theta}_m$, σ_{MOI}, μ_{BI}, and σ_{BI} by maximizing the log of eqn. (7) using the Nelder-Mead optimization algorithm under default settings in the optim function in R[47]. To safeguard against obtaining an estimate that represented a local rather than global optimum, we repeated this optimization procedure 1,000 times under different initial conditions. The initial conditions for each of these runs came from separate draws from a posterior distribution obtained through SMC estimation using the BayesianTools R library[46]. Of the 1,000 estimates of $\vec{\theta}_m$, σ_{MOI}, μ_{BI}, and σ_{BI} that this yielded, we chose the one with the highest log likelihood to derive our maximum-likelihood estimate of $m(t)$ for use in the epidemiological model (Fig S8).

Epidemiological likelihood

The probability of the local incidence data, $\vec{I}$, depends on a total of 51 parameters in addition to $m(t)$ that define $\beta(t)$, with nine for $\vec{\theta}_{s_m}$, nine for $\vec{\theta}_{s_T}$, and three for each of the eleven years spanned by $\vec{\theta}_0$. Although the transmission model (eqn. 3) is stochastic, it does not readily lend itself to calculation of the probability of $\vec{I}$ as a function of these parameters. Consequently, we used a simulation-based approach to approximate the probability of each daily value of I_t under a given value of the 51 model parameters. To do so, we performed 100 simulations of the entire time series of local incidence across 2005-2015, with each simulation driven by data on imported cases feeding into eqn. (3) for a given $\beta(t)$. As local incidence occurred in these simulations, those cases fed back into generating subsequent cases, again following eqn. (3). Using these simulations, we approximated a probability of the local incidence data by treating the number of local cases on a given day as a beta binomial random variable. This assumes that all residents of Guangzhou are subject to a probability of being infected and detected by surveillance as a locally acquired dengue case on each day, with uncertainty in that probability described by a beta distribution. The parameters of that beta distribution enable calculation of the beta binomial probability and were approximated by Bayesian conjugate distributional relationships as $\alpha_t = 1 + \sum_{i=1}^{100} I_{t,i}$ and $\beta_t = N - \sum_{i=1}^{100} I_{t,i} + 1$ [48], where $N = 14{,}040{,}000$. In summary, 100 values of $I_{t,i}$ simulated for each day in 2005-2015 using a single $\beta(t)$ specified

$$\mathcal{L}\left(\vec{\theta}_{s_m}, \vec{\theta}_{s_T}, \vec{\theta}_0\middle|\vec{I}, \vec{T}, m(t)\right) = \prod_t \mathrm{beta\ binomial}(I_t|\alpha_t, \beta_t, N) \quad (9)$$

as the overall likelihood of the epidemiological model parameters.

Epidemiological parameter priors

Given that numerous studies have investigated relationships between temperature, mosquito density, and DENV transmission, we sought to leverage that information by specifying prior

distributions for epidemiological model parameters. Doing so still permits the data to influence parameter estimates in the posterior via the likelihood, but it does so in such a way that parameter values in the posterior are penalized if they deviate too strongly from prior understanding of which parameter values are plausible based on previous work. For $\vec{\theta}_{s_m}$ and $\vec{\theta}_{s_T}$, we used relationships between T, m, and R_0 (which is similar to our transmission coefficient β[41]) recently described by Siraj et al.[8]. In doing so, we assumed that relationships among these variables were identical at all lags τ, given a lack of specific prior understanding of how these relationships vary at different lags. Given that the scales of m and that of Siraj et al.[8] are not directly comparable, we parameterized the prior distribution around values of m with relevance to the time series of $m(t)$ in Guangzhou. That is, at the temperature optimum of 33.3 °C estimated by Siraj et al.[8], we set our prior for β such that $\beta = 0$ when $m = 0$ and $\beta = 2.5$ when $m = 3$. The latter value of m is just above the maximum value estimated for Guangzhou, and the corresponding value of β is equal to the median seasonal estimate of daily R_0 in Iquitos, Peru, a dengue-endemic setting with empirical estimates of seasonal ${R_0}$[49] that Guangzhou should be unlikely to exceed. At the same time, posterior estimates of the parameters did have the flexibility to yield values of β in excess of 2.5 should the data support such values via the likelihood. Consistent with standard theory for mosquito-borne disease epidemiology[50], values of the prior at other temperatures were obtained by reducing the value of β linearly in proportion to m and by the proportion of R_0 from Siraj et al.[8] for other temperatures relative to its value at 33.3 °C. Using 1,000 Monte Carlo samples of the relationship between T, m, and R_0 from Siraj et al.[8], we obtained 1,000 estimates of $\vec{\theta}_{s_m}$ and $\vec{\theta}_{s_T}$ by using the optim function in R to minimize the sum of squared differences between R_0 values from Siraj et al.[8] and corresponding values of β defined by $\vec{\theta}_{s_m}$ and $\vec{\theta}_{s_T}$and with $\beta_0 = 0$. A multivariate normal distribution fitted to those 1,000 estimates of $\vec{\theta}_{s_m}$ and $\vec{\theta}_{s_T}$ represented our prior distribution of those parameters. Separately, we defined the prior distribution of each parameter in $\vec{\theta}_0$ as normally distributed with mean 0 and standard deviation 5, given our expectation that residual variation in $\beta(t)$ not attributable to temperature or mosquito density should be minimal, on average.

Epidemiological model fitting

We obtained an estimate of the posterior distribution of epidemiological parameters using an SMC algorithm implemented in the BayesianTools R library[51]. To assess convergence, we performed three independent runs of the SMC algorithm set to ten iterations of 10,000 samples each. We then calculated the Gelman-Rubin convergence diagnostic statistic across the three independent runs, along with the multivariate potential scale reduction factor (Table S2)[52].

Simulation experiments

Checking model consistency with data

To verify that the behavior of the transmission model was consistent with the data to which it was fitted, we simulated an ensemble of 2,000 realizations of daily local incidence using parameter values drawn from the estimated posterior distribution. These simulations were performed for all of 2005-2015 in the same manner in which the likelihood was approximated; i.e., driven by imported case data and with local transmission following eqns. (1)-(4). We compared simulated and empirical local incidence patterns in two ways. First, we computed Pearson's correlation coefficient between daily local incidence data and median values from the simulation ensemble. Second, we compared simulated and empirical patterns on an annual basis in terms of four features of local incidence patterns: annual incidence, peak weekly incidence, total number of weeks with non-zero local incidence, and number of weeks between the first and last local case. Consistency between simulated and empirical values of these

quantities was assessed using Bayesian p-values, with values > 0.025 and < 0.975 indicating consistency between empirical values and the model-derived ensemble[48].

Factorial simulation experiment

To partition inter-annual variation in local incidence into portions attributable to inter-annual variation in local conditions or importation patterns, we performed a simulation experiment with a two-way factorial design. In this experiment, we grouped temperature, mosquito density, and residual variation in local conditions together as one set of predictor variables and importation patterns as the other. Each year from 2005 to 2015 was considered as a factor for each set of predictors. An ensemble of 1,000 simulations was generated for each of the 122 combinations of 11 years of each of the two sets of predictors. For example, with temperatures, mosquito densities, and $\beta_0(t)$ values from 2005, 1,000 simulations were performed with imported cases from each of 2005-2015, and likewise for temperatures, mosquito densities, and $\beta_0(t)$ values from 2006-2015. We summed annual local incidence for each of these 122,000 simulations and performed a two-way analysis of variance, resulting in estimates of the variation (defined in terms of sum of squared error, SSQ) in annual incidence attributable to local conditions, to importation, and to a portion unexplained by either predictor set due to the stochastic nature of the simulations.

To quantify the overall portion of variation attributable to each predictor variable, we performed an additional simulation experiment with a four-way factorial design. In this experiment, we interchanged temperature, mosquito density, $\beta_0(t)$ values, and importation patterns from different years, again considering each year as a factor for each predictor variable. An ensemble of 1,000 simulations was generated for each of the 14,641 combinations of 11 years of all four predictors. Similar to the two-way factorial experiment, we summed annual local incidence for each of these simulations and performed a four-way analysis of variance. This resulted in direct estimates of the variation in annual incidence attributable to temperature, mosquito density, $\beta_0(t)$ values, importation patterns, and to a portion attributable to stochasticity.

ACKNOWLEDGMENTS

This study was funded by grants from the National Science Fund for Distinguished Young Scholars (81525023) (HY), National Natural Science Foundation of China (81663498) (HY), University of Notre Dame (Arthur J. Schmitt Leadership Fellowships in Science and Engineering) (RO), National Science Foundation (1641130) (AP), Defense Advanced Research Projects Agency (D16AP00114) (AP), Bill and Melinda Gates Foundation (OPP1106427, OPP1032350, OPP1134076, OPP1094793) (AT), Department for International Development (AT), and Clinton Health Access Initiative (AT). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Patz, J. ., Campbell-Lendrum, D., Holloway, T. & Foley, J. A. Impact of regional climate change on human health. *Nature* **438,** 310–317 (2005).
2. Lafferty, K. D. The ecology of climate change and infectious diseases. *Ecology* **90,** 888–900 (2009).
3. Pascual, M., Rodo, X., Ellner, S. P., Colwell, R. & Bouma, M. J. Cholera Dynamics and El Nino–Southern Oscillation. *Science.* **289,** 1766–1769 (2000).
4. Moore, S. M. *et al.* El Niño and the shifting geography of cholera in Africa. *Proc. Natl. Acad. Sci.* **114,** 4436–4441 (2017).
5. Caminade, C. *et al.* Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proc. Natl. Acad. Sci.* **114,** 119–124 (2017).

6. Lowe, R. *et al.* Climate services for health: predicting the evolution of the 2016 dengue season in Machala, Ecuador. *Lancet Planet. Heal.* **1,** e142–e151 (2017).
7. Ryan, S. J. *et al.* Mapping physiological suitability limits for malaria in Africa under climate change. *Vector borne zoonotic Dis.* **15,** 718–25 (2015).
8. Siraj, A. S. *et al.* Temperature modulates dengue virus epidemic growth rates through its effects on reproduction numbers and generation intervals. *PLoS Negl. Trop. Dis.* 1–19 (2017).
9. Johansson, M. A., Reich, N. G., Hota, A., Brownstein, J. S. & Santillana, M. Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Sci. Rep.* **6,** 1–11 (2016).
10. Morin, C. W., Comrie, A. C. & Ernst, K. Climate and Dengue Transmission: Evidence and Implications. *Environ. Health Perspect.* **121,** 1264–1272 (2013).
11. Koelle, K. & Pascual, M. Disentangling Extrinsic from Intrinsic Factors in Disease Dynamics : A Nonlinear Time Series Approach with an Application to Cholera. *Am. Nat.* **163,** 901–913 (2004).
12. Koelle, K., Rodó, X., Pascual, M., Yunus, M. & Mostafa, G. Refractory periods and climate forcing in cholera dynamics. *Nature* **436,** 696–700 (2005).
13. Wearing, H. J. & Rohani, P. Ecological and immunological determinants of dengue epidemics. *Proc Natl Acad Sci U S A* **103,** 11802–11807 (2006).
14. Cummings, D. A. T. *et al.* The impact of the demographic transition on dengue in Thailand: Insights from a statistical analysis and mathematical modeling. *PLoS Med.* **6,** 1–10 (2009).
15. Grubaugh, N. D. *et al.* Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature* **546,** 401–405 (2017).
16. Lai, S. *et al.* The changing epidemiology of dengue in China, 1990-2014: a descriptive analysis of 25 years of nationwide surveillance data. *BMC Med.* **13,** 100 (2015).
17. van Panhuis, W. G. *et al.* Region-wide synchrony and traveling waves of dengue across eight countries in Southeast Asia. *Proc. Natl. Acad. Sci.* **112,** 13069–13074 (2015).
18. Sang, S. *et al.* Predicting Unprecedented Dengue Outbreak Using Imported Cases and Climatic Factors in Guangzhou, 2014. *PLoS Negl. Trop. Dis.* **9,** 1–12 (2015).
19. Xu, L. *et al.* Climate variation drives dengue dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 113–118 (2016).
20. Cheng, Q. *et al.* Climate and the Timing of Imported Cases as Determinants of the Dengue Outbreak in Guangzhou, 2014: Evidence from a Mathematical Model. *PLoS Negl. Trop. Dis.* **10,** 1–22 (2016).
21. Cheng, Q. *et al.* The interplay of climate, intervention and imported cases as determinants of the 2014 dengue outbreak in Guangzhou. *PLoS Negl. Trop. Dis.* **11,** 1–24 (2017).
22. Li, M.-T. *et al.* The Driving Force for 2014 Dengue Outbreak in Guangdong, China. *PLoS One* **11,** e0166211 (2016).
23. Cao, Z. *et al.* Individual and interactive effects of socio-ecological factors on dengue fever at fine spatial scale: A geographical detector-based analysis. *Int. J. Environ. Res. Public Health* **14,** (2017).
24. Zhu, G., Liu, J., Tan, Q. & Shi, B. Inferring the Spatio-temporal Patterns of Dengue Transmission from Surveillance Data in Guangzhou, China. *PLoS Negl. Trop. Dis.* **10,** 1–20 (2016).
25. Finkenstädt, B. F. & Grenfell, B. T. Time series modelling of childhood diseases: a dynamical systems approach. *Appl. Stat.* **49,** 187–205 (2000).
26. White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C. & White, C. Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos* **123,** 385–388 (2014).

27. Clapham, H. E., Cummings, D. A. T. & Johansson, M. A. Immune status alters the probability of apparent illness due to dengue virus infection : evidence from a pooled analysis across multiple cohort and cluster studies. **57,** 1–12 (2017).
28. Ferguson, N. M. *et al.* Modeling the impact on virus transmission of Wolbachia-mediated blocking of dengue virus infection of Aedes aegypti. *Sci. Transl. Med.* **7,** (2015).
29. Fried, J. R. *et al.* Serotype-specific differences in the risk of dengue hemorrhagic fever: An analysis of data collected in Bangkok, Thailand from 1994 to 2006. *PLoS Negl. Trop. Dis.* **4,** 1–6 (2010).
30. Rico-Hesse, R. Microevolution and virulence of dengue viruses. *Adv Virus Res* **59,** 315–341 (2003).
31. Reiner, R. C. *et al.* Estimating malaria transmission from humans to mosquitoes in a noisy landscape. *J. R. Soc. Interface* **12,** 20150478 (2015).
32. Ernst, K. C. *et al.* Aedes aegypti (Diptera: Culicidae) Longevity and Differential Emergence of Dengue Fever in Two Cities in Sonora, Mexico. *J. Med. Entomol.* **54,** 204–211 (2017).
33. Li, Z. *et al.* Dengue Baidu Search Index data can improve the prediction of local dengue epidemic: A case study in Guangzhou, China. *PLoS Negl. Trop. Dis.* **11,** 1–13 (2017).
34. Ye, X., Li, S., Yang, X. & Qin, C. Use of Social Media for the Detection and Analysis of Infectious Diseases in China. *ISPRS Int. J. Geo-Information* **5,** 156 (2016).
35. Funk, S. *et al.* Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics* **10,** 21–25 (2015).
36. Metcalf, C. J. E. *et al.* Identifying climate drivers of infectious disease dynamics: Recent advances and challenges ahead. *Proc. R. Soc. B Biol. Sci.* **284,** (2017).
37. Cauchemez, S. & Ferguson, N. M. Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *J. R. Soc. Interface* **5,** 885–897 (2008).
38. Guo, R. N. *et al.* The prevalence and endemic nature of dengue infections in Guangdong, South China: An epidemiological, serological, and etiological study from 2005-2011. *PLoS One* **9,** 1–8 (2014).
39. Qiaoli, Z. *et al.* Maiden outbreak of chikungunya in Dongguan city, Guangdong province, China: Epidemiological characteristics. *PLoS One* **7,** 1–7 (2012).
40. Vector surveillance.
41. Perkins, T. A., Metcalf, C. J. E., Grenfell, B. T. & Tatem, A. J. Estimating Drivers of Autochthonous Transmission of Chikungunya Virus in its Invasion of the Americas. *PLoS Curr.* **7,** (2015).
42. Xia, Y., Bjørnstad, O. N. & Grenfell, B. T. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *Am. Nat.* **164,** 267–81 (2004).
43. Grenfell, B. T., Bjørnstad, O. N. & Finkenstädt, B. F. Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the TSIR model. *Ecol. Monogr.* **72,** 185–202 (2002).
44. Perkins, T. A. *et al.* A review of transmission models of dengue: a quantitative and qualitative analysis of model feature. in *Dengue and Dengue Hemorrhagic Fever* (eds. Gubler, D. J., Ooi, E. E. & Farrar, J.) (CABI Publishing, 2014).
45. Ramsay, J. O., Wickham, H., Graves, S. & Hooker, G. fda: Functional Data Analysis. *R Packag. version 2.4.4.* (2014).
46. Hartig, F., Minunno, F. & Paul, S. BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics. (2017).
47. R Development Core Team. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna Austria* **0,** {ISBN} 3-900051-07-0 (2016).
48. Hobbs, N. T. & Hooten, M. B. *Bayesian Models: A Statistical Primer for Ecologists*. (Princeton University Press, 2015).

49. Reiner, R. C. *et al.* Time-varying, serotype-specific force of infection of dengue virus. *Proc. Natl. Acad. Sci.* **111,** E2694–E2702 (2014).
50. Smith, D. L. & Ellis McKenzie, F. Statics and dynamics of malaria infection in Anopheles mosquitoes. *Malar. J.* **3,** 13 (2004).
51. Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T. & Huth, A. Statistical inference for stochastic simulation models - theory and application. *Ecol. Lett.* **14,** 816–827 (2011).
52. Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. *Bayesian Data Analysis*. *Chapman Texts in Statistical Science Series* (2004). doi:10.1007/s13398-014-0173-7.2

Supporting Information for

Inter-annual variation in seasonal dengue epidemics driven by multiple interacting factors in Guangzhou, China

Rachel J. Oidtman[1], Shengjie Lai[2,3,4], Zhoujie Huang[2], Amir S. Siraj[1], Robert C. Reiner[5], Andrew J. Tatem[3,4], T. Alex Perkins*[1], Hongjie Yu*[2]

[1] Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA
[2] School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China
[3] WorldPop, Department of Geography and Environment, University of Southampton, Southampton, UK
[4] Flowminder Foundation, Stockholm, Sweden
[5] Institute for Health and Metrics and Evaluation, University of Washington, Seattle, WA, USA

* Corresponding authors: yhj@fudan.edu.cn and taperkins@nd.edu

SUPPLEMENTARY TABLES

Paper	Location	Time step	Time frame	Type of model	Factors / covariates included	Main conclusion
[1]	Guangzhou	Monthly	2005-2015	Statistical	Avg. temperature, avg. min. temperature, avg. max. temperature, accumulated precipitation, number of days with rainfall, MOI, and BI	Climate conditions explain temporal dynamics of dengue incidence
[2]	Guangzhou	Monthly	2006-Sept. 2014	Statistical	Local min. temperature, accumulative precipitation	Minimum temperature in previous month and accumulative precipitation with 3-month lag can project dengue outbreaks of 2013 and 2014
[3]	Guangdong	Weekly	2014	Mechanistic	Weekly avg. temperature, and weekly precipitation	Delayed mosquito control, continuous importations from April-July, transmission of asymptomatic infections, and high precipitation from May-August are causal factors for unprecedented outbreak
[4]	Guangzhou	Daily	2013-2014	Mechanistic, deterministic	Temperature, rainfall, and evaporation	In 2013 and 2014, date of first imported case and unusually high precipitation, and in 2014 only, delayed interventions and vertical transmission, are factors responsible for patterns of moderate outbreak in 2013 and much larger outbreak in 2014
[5]	Guangzhou	Daily	2014	Spatial	Urbanization level, ratio of urban village, road density, population density, GDP, NDVI, temperature, and precipitation	Temperature, precipitation, road density, and water body area were the dominant factors of dengue fever in 2014 outbreak
[6]	Guangzhou	Daily	Sept. 2014-Nov. 2014	Mechanistic	Population density, human mobility (transportation), temperature, rainfall, humidity, MOI, and BI	Urbanization, vector activities, and human behavior played significant roles in shaping the 2014 dengue outbreak and the patterns of its spread
[7]	Guangzhou	Daily	2013-2014	Mechanistic, stochastic	Tourist exchange, temperature, precipitation, BI, and MOI	Higher number of imported cases in May and June were the most important determinants of dengue outbreaks

Table S1. Model attributes, covariates included, and conclusions of papers that investigated dengue incidence in Guangzhou or Guangdong in 2014 and other years.

Parameter	Point estimate	Upper C.I.
$\theta_{s_{m_1}}$	1.03	1.08
$\theta_{s_{m_2}}$	1.03	1.09
$\theta_{s_{m_3}}$	1.03	1.10
$\theta_{s_{m_4}}$	1.01	1.04
$\theta_{s_{m_5}}$	1.03	1.11
$\theta_{s_{m_6}}$	1.03	1.08
$\theta_{s_{m_7}}$	1.00	1.01
$\theta_{s_{m_8}}$	1.00	1.00
$\theta_{s_{m_9}}$	1.00	1.00
$\theta_{s_{T_1}}$	1.04	1.10
$\theta_{s_{T_2}}$	1.03	1.09
$\theta_{s_{T_3}}$	1.03	1.09
$\theta_{s_{T_4}}$	1.04	1.12
$\theta_{s_{T_5}}$	1.02	1.07
$\theta_{s_{T_6}}$	1.02	1.07
$\theta_{s_{T_7}}$	1.02	1.06
$\theta_{s_{T_8}}$	1.00	1.01
$\theta_{s_{T_9}}$	1.00	1.01
θ_{0_1}	1.05	1.15
θ_{0_2}	1.01	1.01
θ_{0_3}	1.15	1.45
θ_{0_4}	1.03	1.07
θ_{0_5}	1.05	1.18
θ_{0_6}	1.06	1.19
θ_{0_7}	1.04	1.14
θ_{0_8}	1.04	1.12
θ_{0_9}	1.09	1.28
$\theta_{0_{10}}$	1.06	1.17
$\theta_{0_{11}}$	1.03	1.08
$\theta_{0_{12}}$	1.01	1.02
$\theta_{0_{13}}$	1.05	1.16
$\theta_{0_{14}}$	1.00	1.01
$\theta_{0_{15}}$	1.02	1.06
$\theta_{0_{16}}$	1.19	1.56
$\theta_{0_{17}}$	1.04	1.05
$\theta_{0_{18}}$	1.09	1.18
$\theta_{0_{19}}$	1.04	1.12
$\theta_{0_{20}}$	1.03	1.08
$\theta_{0_{21}}$	1.06	1.17
$\theta_{0_{22}}$	1.09	1.21

$\theta_{0_{23}}$	1.03	1.07
$\theta_{0_{24}}$	1.04	1.08
$\theta_{0_{25}}$	1.04	1.14
$\theta_{0_{26}}$	1.04	1.09
$\theta_{0_{27}}$	1.02	1.07
$\theta_{0_{28}}$	1.03	1.05
$\theta_{0_{29}}$	1.05	1.17
$\theta_{0_{30}}$	1.03	1.06
$\theta_{0_{31}}$	1.10	1.31
$\theta_{0_{32}}$	1.03	1.10
$\theta_{0_{33}}$	1.16	1.46

Table S2. Gelman-Rubin convergence diagnostic based on three independent SMC sampling routines. Convergence is diagnosed to have occurred when the upper confidence interval is close to 1. The multivariate potential scale reduction factor is 1.72.

SUPPLEMENTARY FIGURES

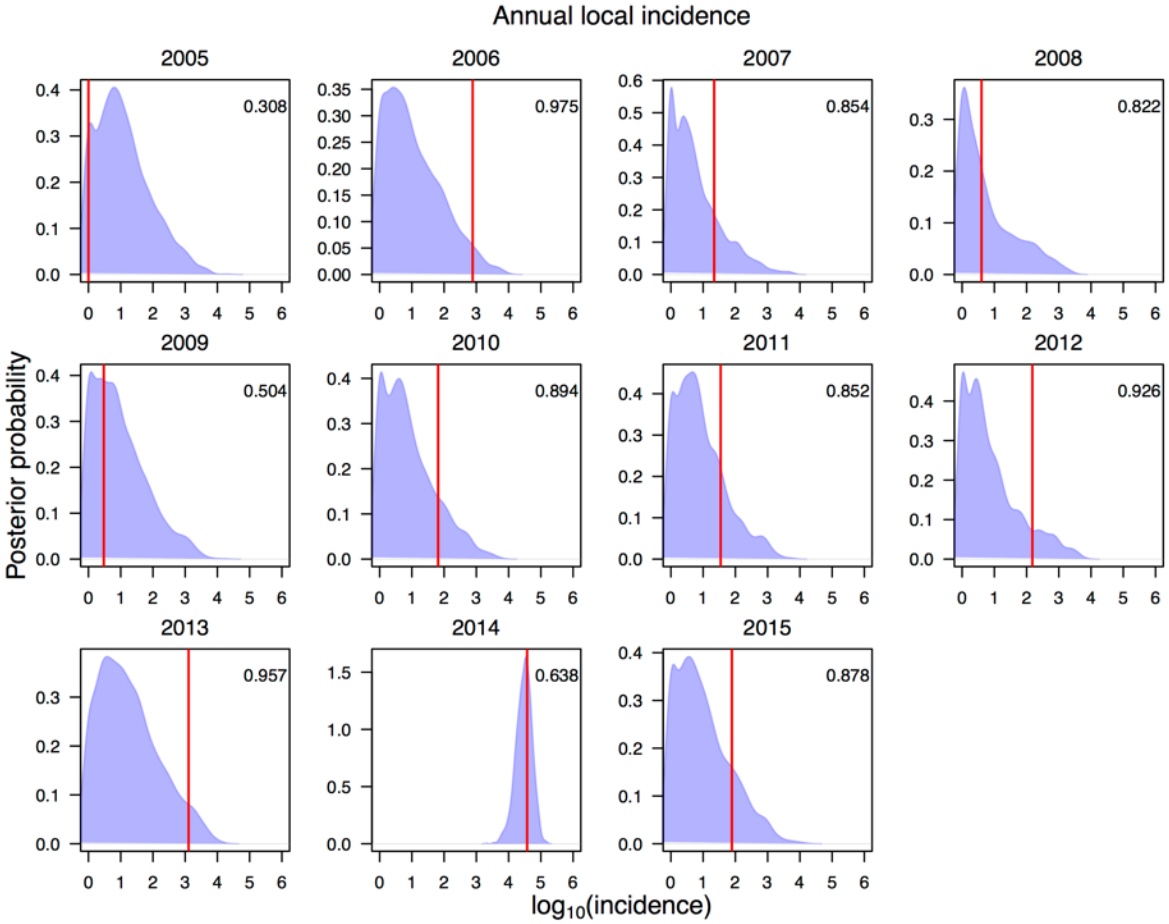

Fig S1. Posterior predictive distribution of total dengue incidence for each year based on 2,000 simulations from the fitted transmission model. The red line shows observed annual local incidence for each year. The number in each panel indicates the Bayesian p-value, with values between 0.025 and 0.975 indicating model consistency with the data.

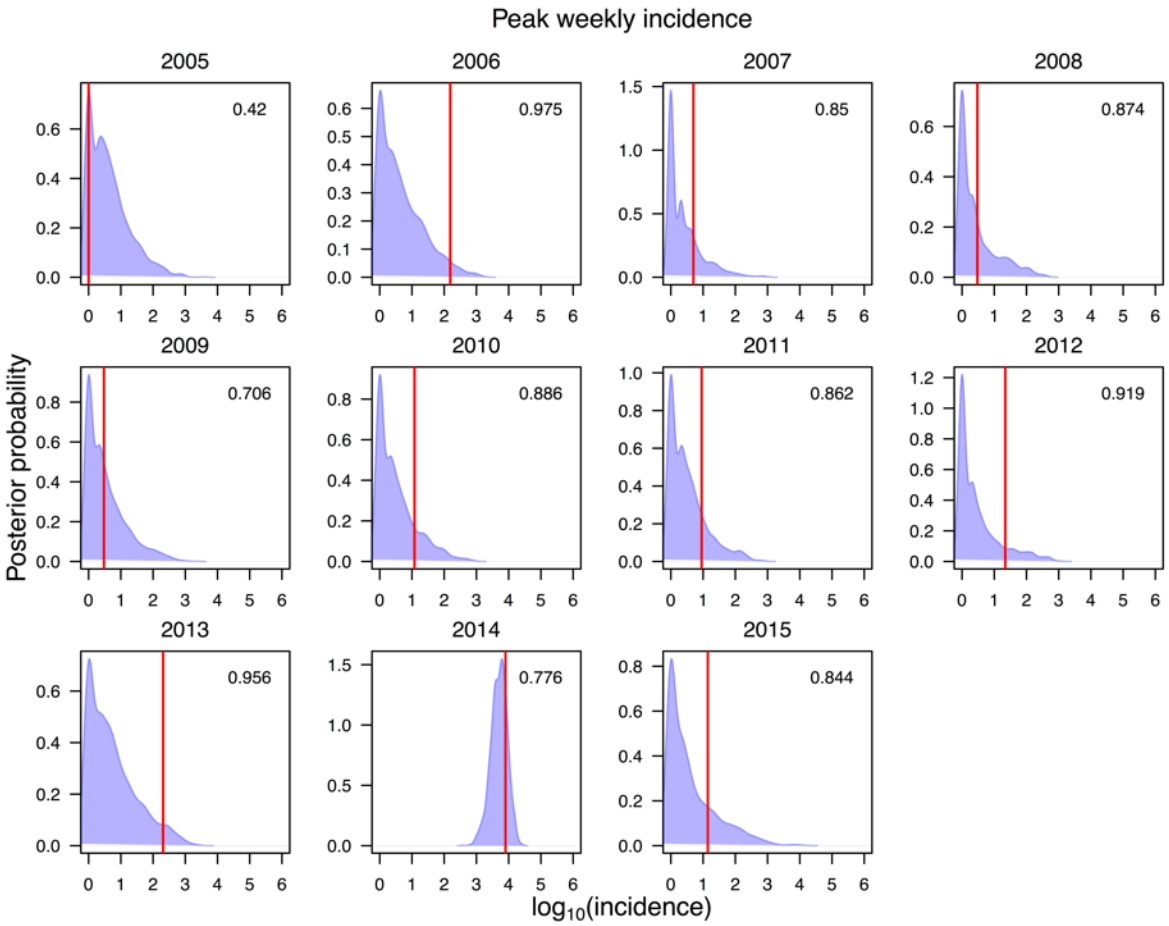

Fig S2. Posterior predictive distribution of peak weekly incidence for each year based on 2,000 simulations from the fitted transmission model. The red line shows observed peak weekly dengue incidence for each year. The number in each panel indicates the Bayesian p-value, with values between 0.025 and 0.975 indicating model consistency with the data.

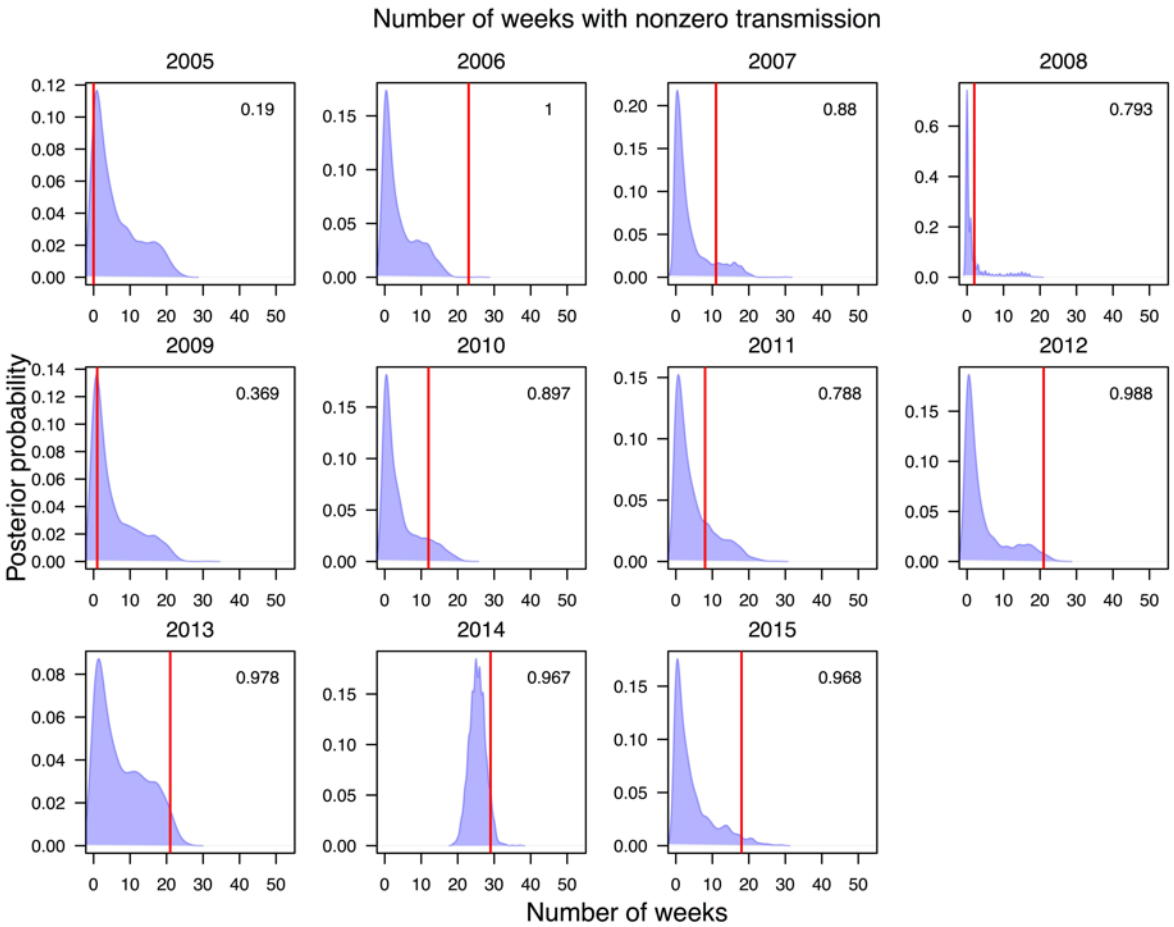

Fig S3. Posterior predictive distribution of the number of weeks with non-zero dengue incidence in each year based on 2,000 simulations from the fitted transmission model. The red line shows observed number of weeks with non-zero dengue incidence for each year. The number in each panel indicates the Bayesian p-value, with values between 0.025 and 0.975 indicating model consistency with the data.

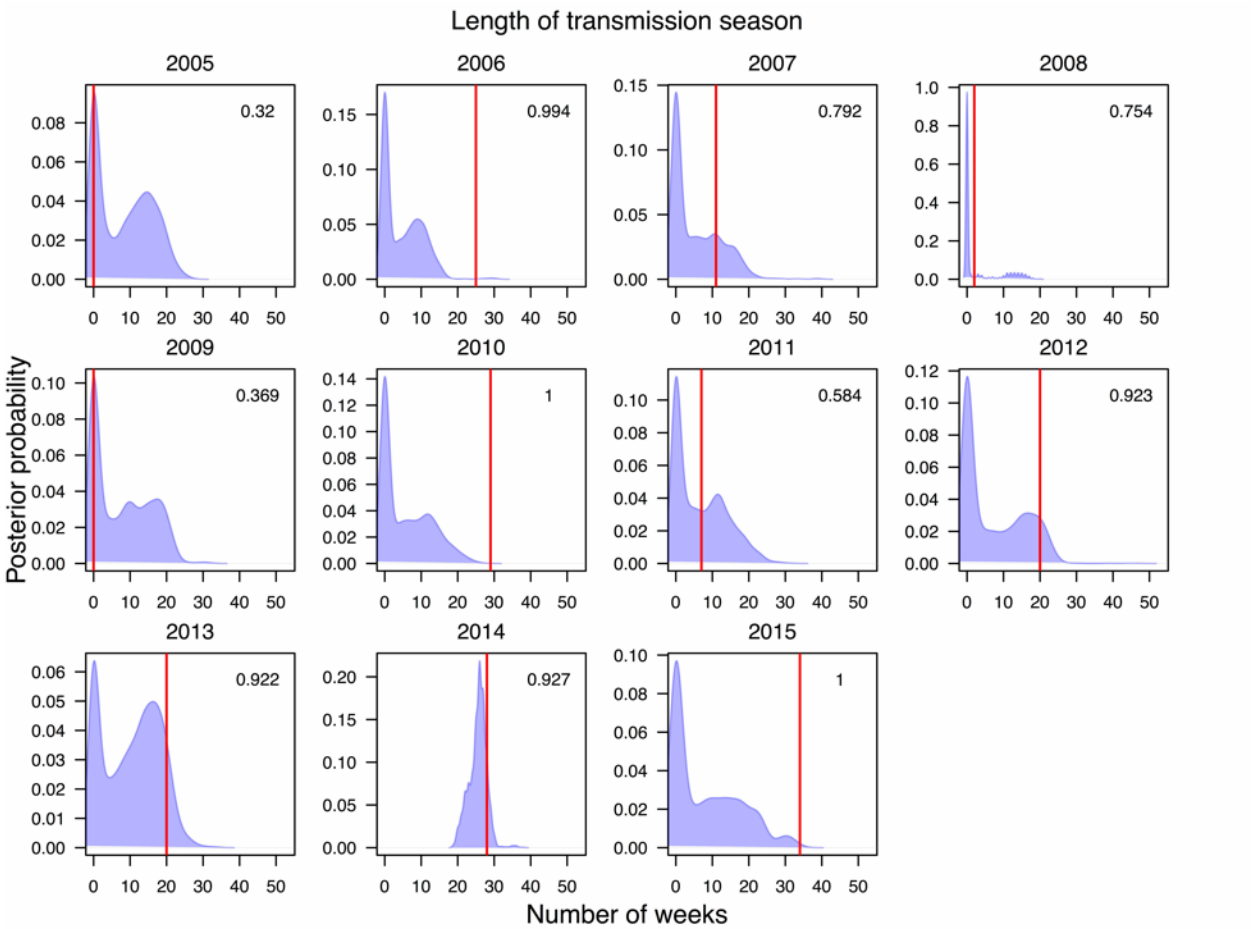

Fig S4. Posterior predictive distribution of the difference between the first non-zero dengue incidence week to the last non-zero dengue incidence week in each year based on 2,000 simulations from the fitted transmission model. The red line shows observed difference between the first non-zero dengue incidence week to the last non-zero dengue incidence day for each year. The number in each panel indicates the Bayesian p-value, with values between 0.025 and 0.975 indicating model consistency with the data.

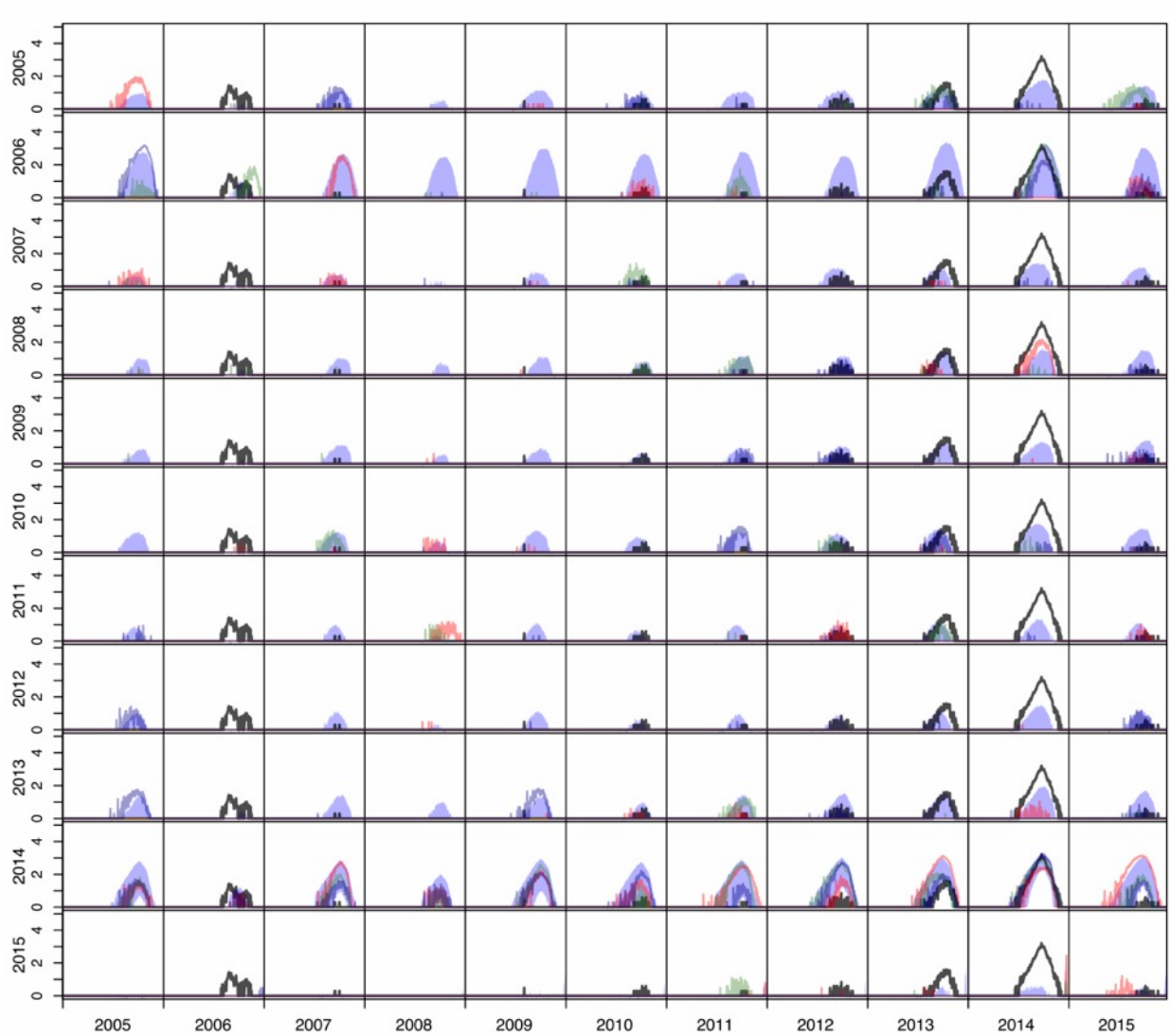

Fig S5. Epidemic trajectories ($\log_{10}$ scale) from the factorial experiment for each combination of local condition years (rows) and imported case years (columns). The shaded regions represent 95% posterior predictive intervals. Black lines represent observed values of daily local incidence. Colored lines show three different random samples from the epidemic trajectory posterior.

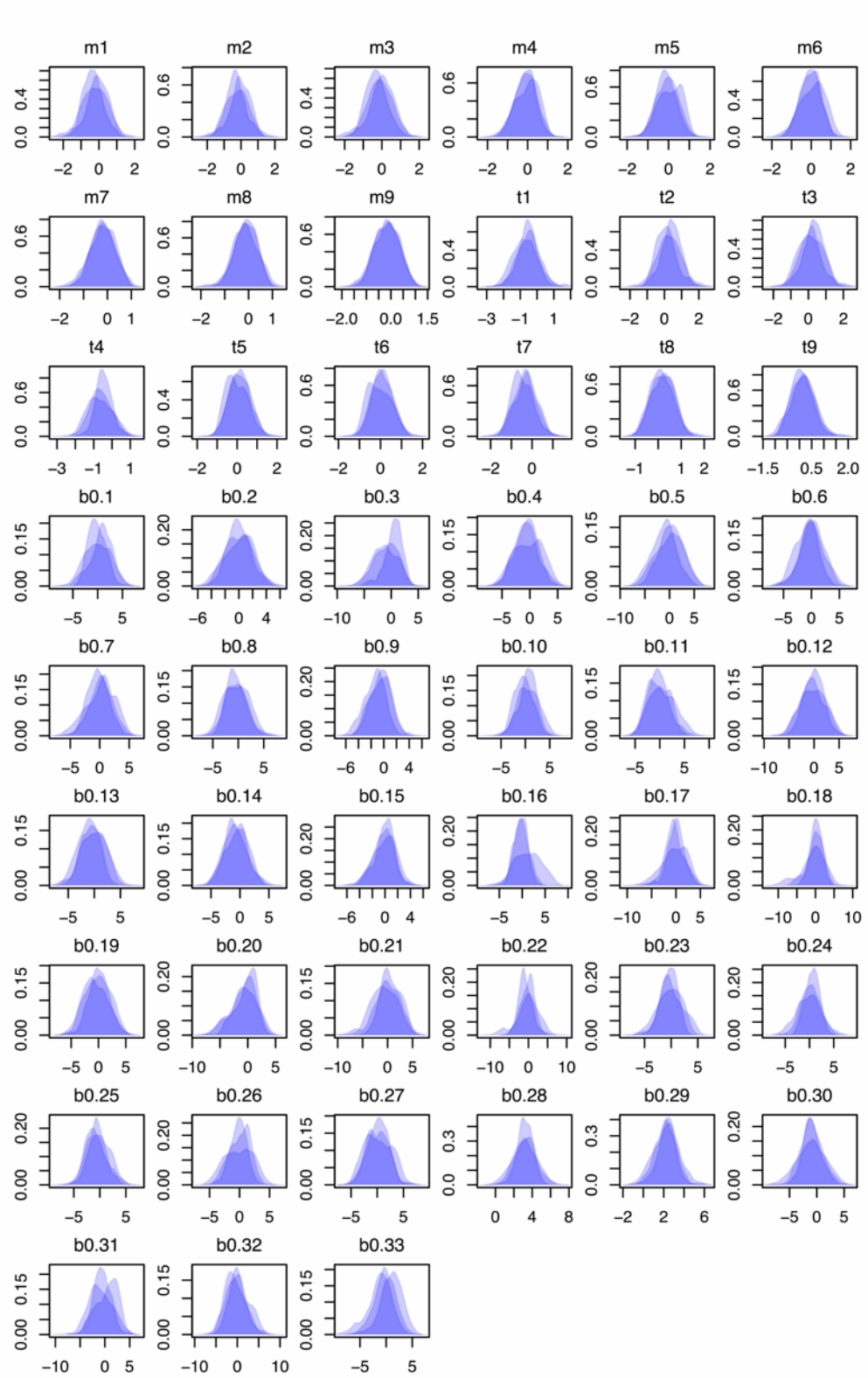

Fig S6. Posterior distributions for parameters estimated using Sequential Monte Carlo sampling. Within each plot, each shaded region represents the posterior distribution for a parameter from one independent SMC run.

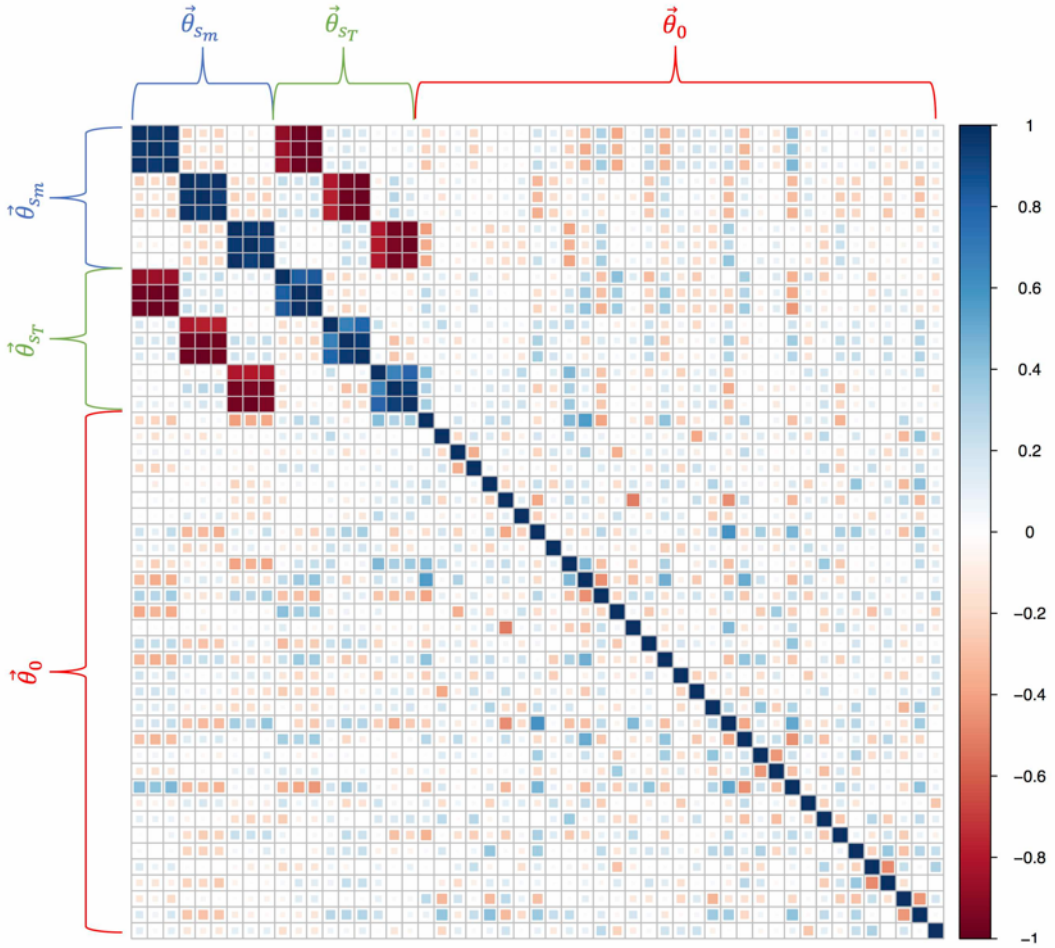

Fig S7. Correlation matrix for posterior estimates of parameters. Brackets indicate which entries in the matrix belong to which set of parameters. Dark blue corresponds to a large positive correlation and dark red corresponds to a large negative correlation.

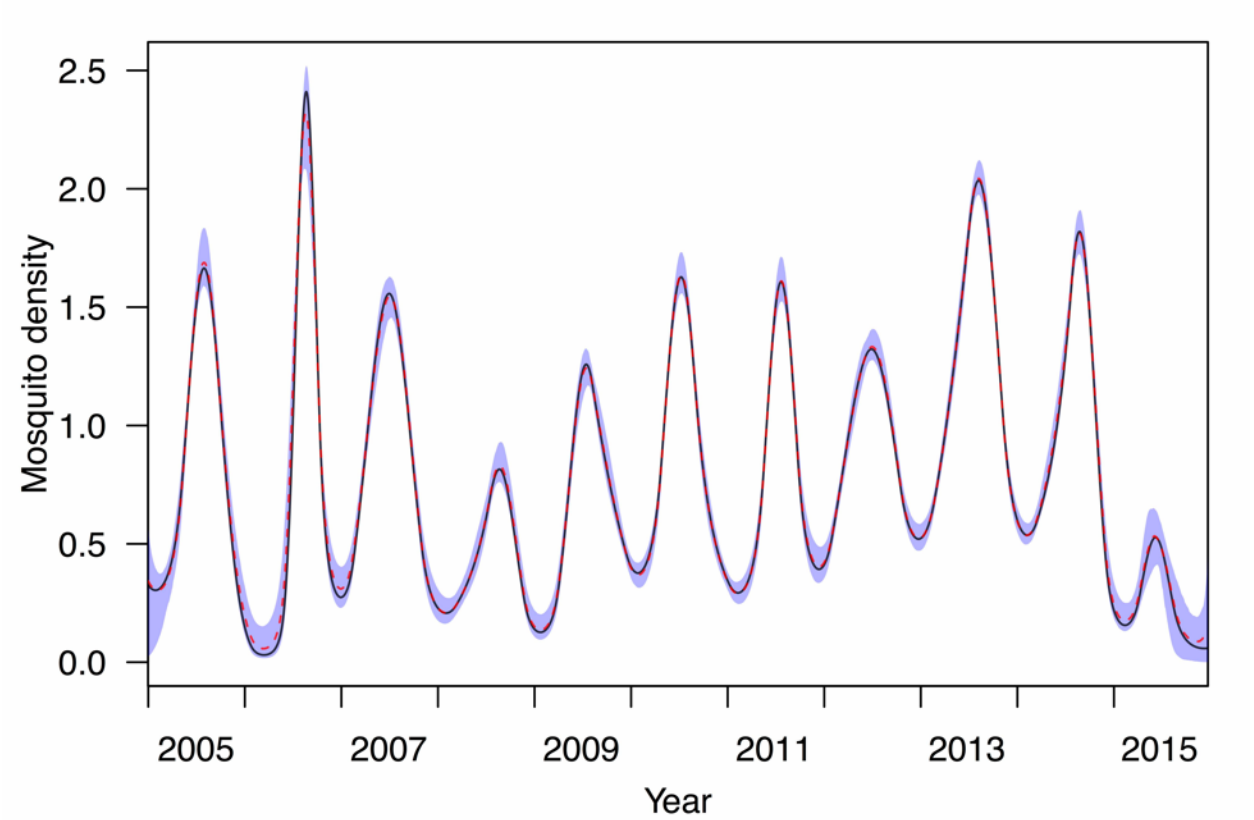

Fig S8. Confidence interval surrounding mosquito spline estimates. Blue shaded region is the 95% confidence interval. Red dotted line is the median mosquito spline estimate. Black line is the maximum likelihood estimate.

References

1. Xu L, Stige LC, Chan K-S, Zhou J, Yang J, Sang S, et al. Climate variation drives dengue dynamics. Proc Natl Acad Sci. 2016;114(1):201618558.
2. Sang S, Gu S, Bi P, Yang W, Yang Z, Xu L, et al. Predicting Unprecedented Dengue Outbreak Using Imported Cases and Climatic Factors in Guangzhou, 2014. PLoS Negl Trop Dis. 2015;9(5):1–12.
3. Li M-T, Sun G-Q, Yakob L, Zhu H-P, Jin Z, Zhang W-Y. The Driving Force for 2014 Dengue Outbreak in Guangdong, China. PLoS One. 2016;11(11):e0166211.
4. Cheng Q, Jing Q, Spear RC, Marshall JM, Yang Z, Gong P. Climate and the Timing of Imported Cases as Determinants of the Dengue Outbreak in Guangzhou, 2014: Evidence from a Mathematical Model. PLoS Negl Trop Dis. 2016;10(2):1–22.
5. Cao Z, Liu T, Li X, Wang J, Lin H, Chen L, et al. Individual and Interactive Effects of Socio-Ecological Factors on Dengue Fever at Fine Spatial Scale: A Geographical Detector-Based Analysis. Int J Environ Res Public Health. 2017;14(7):795.
6. Zhu G, Liu J, Tan Q, Shi B. Inferring the Spatio-temporal Patterns of Dengue Transmission from Surveillance Data in Guangzhou, China. PLoS Negl Trop Dis. 2016;10(4):1–20.
7. Cheng Q, Jing Q, Spear RC, Marshall JM, Yang Z, Gong P. The interplay of climate, intervention and imported cases as determinants of the 2014 dengue outbreak in Guangzhou. PLoS Negl Trop Dis. 2017;11(6):e0005701.