
Randomness of Dengue Outbreaks on the Equator

Yirong Chen, Alex R. Cook, Alisa X.L. Lim

A simple mathematical model without seasonality indicated that the apparently chaotic dengue epidemics in Singapore have characteristics similar to epidemics resulting from chance. Randomness as a sufficient condition for patterns of dengue epidemics in equatorial regions calls into question existing explanations for dengue outbreaks there.

Dengue, a vectorborne infectious disease, has complex epidemiologic dynamics (1). The recent expansion of the range of dengue makes this disease a considerable public health concern worldwide (2). In the city-state of Singapore, the number of dengue cases has increased dramatically since the 1990s, and all 4 serotypes of the dengue virus are endemic (3). Cyclical outbreaks of dengue of increasing magnitude have been observed with a cycle of 5–6 years (4), but this pattern appeared to cease in 2005, and no obvious cycle has occurred since then. Although other tropical and subtropical countries in Southeast Asia have distinct seasonality (5) so that dengue epidemics occur at distinct and predictable times of the year (6), Singapore's proximity to the equator gives it an aseasonal climate, and the timing of dengue epidemics is irregular (7,8).

Many factors have been postulated to contribute to dengue's spread in Singapore, such as a consistently warm and humid climate that favors year-round vector proliferation, high urbanization, and a tendency for vectors to live in human residences (9). The extent to which these factors affect dengue epidemics in aseasonal Singapore, if they do at all, is unclear. Competing explanations for the timing of large dengue outbreaks in Singapore can be found in the literature. One study attributes dengue epidemics to conducive temperatures and precipitation variations (10); another attributes them to variable maximum and minimum temperatures (11). Rainfall and temperature have been shown to be related to dengue outbreaks in Brazil, another equatorial country (12).

The tendency to see patterns where none exists has been well recognized. When 2 events happen contemporarily and a plausible story connects the events, the tendency to assume that 1 causes the other is strong (13). Cancer cases

cluster around mobile phone masts (base stations), not because the radiation from a mast is carcinogenic at typical exposures but because numerous masts exist and occasionally cancer cases cluster together, similarly to spilled grains of rice (14). A study in the heuristics and biases program discusses a famous example from sports (15), which are notorious for stories being concocted around essentially chance outcomes. Basketball fans, coaches, and pundits often believe that players have "hot hand" streaks when they have a run of good form, making many shots in succession and playing above their usual level during a match. The study systematically deconstructed this belief by a series of statistical tests that showed that the patterns of actual hits and misses was consistent with mere chance—analogue to sequences of coin tosses rather than an illusory hot hand (15).

In probabilistic models, chance is represented by error terms, or noise, encompassing all the many complicating factors that are not worth including in the systematic signal. Past models for dengue in Singapore have accounted for chance alongside systematic effects of the weather and other factors (10,11). However, is chance alone sufficient to explain the frequent, large, and ostensibly chaotic outbreaks we observe? We sought to assess whether the rise and fall of dengue outbreaks from week to week in Singapore come in runs or are indistinguishable from random noise and thereby whether it is necessary to consider other possible drivers of these epidemics.

The Study

We reviewed data on the weekly incidence of clinically diagnosed dengue in Singapore during 2003–2012. We compared the number of dengue cases per week to a simple simulation model (online Technical Appendix, <http://wwwnc.cdc.gov/EID/article/21/9/14-1030-Techapp.pdf>) with no environmental drivers other than the dependence of weekly number of cases from up to 4 weeks before. Summaries of observed incidence and of the simulated aseasonal model were compared for assessing proximity of the behavior of observed cases to the behavior of simulated cases.

The simulation model used was a standard autoregressive time series model in which the number of cases during any week affects the mean number of cases for the 4 weeks that follow. We allowed the simulated number to have a random variation around that mean; data were log-transformed to ensure that incidence was positive. The fitted autoregressive model was used to simulate synthetic dengue outbreaks over multiple decades, and incidence

Author affiliations: National University of Singapore, Singapore (Y. Chen, A.R. Cook, A.X.L. Lim); National University Health System, Singapore (Y. Chen, A.R. Cook); and Yale–NUS College, Singapore (A.R. Cook)

DOI: <http://dx.doi.org/10.3201/eid2109.141030>

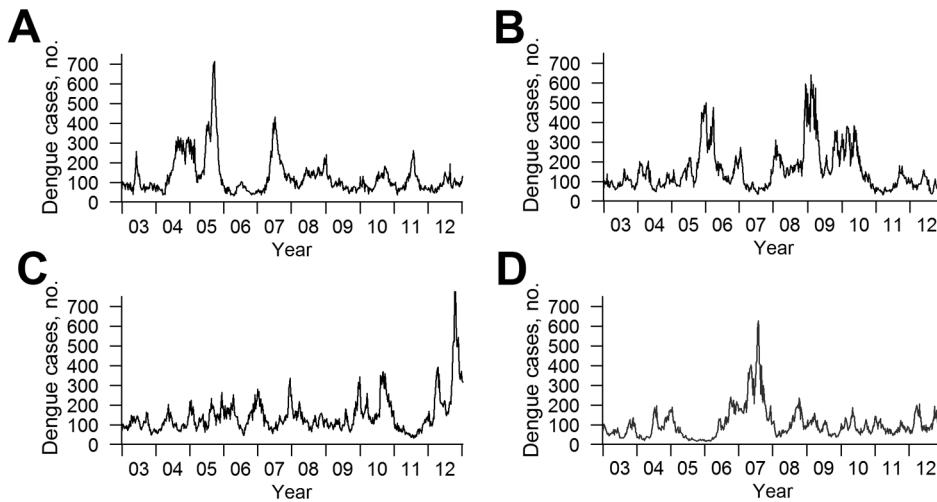


Figure 1. Weekly trends for observed and simulated dengue incidence, 2003–2012, Singapore. A) Weekly trends for the actual scenario of observed dengue incidence. B–D) Three randomly generated simulated scenarios from the aseasonal model described in the text and the online Technical Appendix (<http://wwwnc.cdc.gov/EID/article/21/9/14-1030-Techapp.pdf>). Although the peaks are not synchronized, similar patterns can be discerned; large and small outbreaks of similar scale and frequency occur in all 4 scenarios.

of simulated outbreaks was compared with observed incidence. We devised a series of statistical measures that were inspired by the “hot hand” in basketball study (15) and that might falsify the model that accounted for chance alone. This model included correlation between dengue incidence by week and the preceding week (the autocorrelation function), the probability distribution for the weekly incidence aggregated over 10 years, the distribution of the annual number of cases, the maximum number of cases observed over the previous decade, and the probability of a rise in

incidence each week following a series of rises (i.e., the possible beginning of an epidemic) or a series of declines (i.e., the possible ending of an epidemic). We also created simulated trajectories (Figure 1).

Conclusions

For all metrics considered, the actual scenario (i.e., the observed dengue incidence) was fully consistent with the aseasonal model; both the autocorrelation function (Figure 2, panel A) and the cumulative probability of dengue

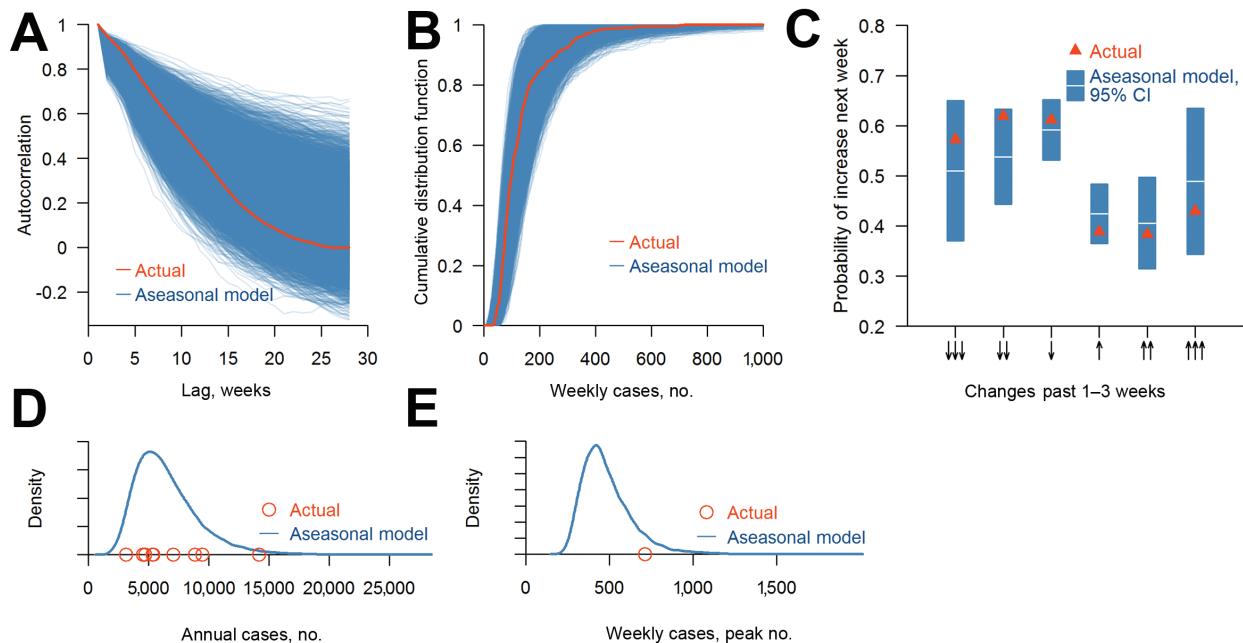


Figure 2. Comparison of observed dengue incidence and incidence from simulated aseasonal models, 2003–2012, Singapore. A) Distribution of actual and simulated autocorrelation functions at different time lags (e.g., this week versus next week; last week versus next week, etc.) B) Distribution of cumulative distribution function of the simulated weekly number of dengue cases and cumulative density function of the actual numbers of cases. C) Conditional probabilities of an increase in number of dengue cases and 95% CIs for simulated scenario and actual data, given 1–3 consecutive decreases or increases. D) Density plot of simulated and actual annual number of dengue cases. E) Density plot of simulated 10-year maximum number of cases and actual 10-year number of cases.

incidence (Figure 2, panel B) from the historical incidence data lie within the distribution resulting from the aseasonal model. The probabilities of an increase in incidence each week that follows a series of rises or falls and corresponding 95% CIs calculated on the basis of simulations from the aseasonal model all include the proportions observed historically (Figure 2, panel C). Furthermore, the distribution of the annual incidence (Figure 2, panel D) and the maximum observed incidence over the decade (Figure 2, panel E) are consistent with the aseasonal model. Similarly, the number of successive increases or decreases over the decade was consistent with chance ($p = 0.18$).

These metrics are not conventional measures of dengue surveillance data; they capture more complex, emergent properties of the epidemic process. However, our findings show that, for dengue incidence in equatorial Singapore, where average monthly temperatures vary only from 26°C–28°C, randomness alone is sufficient to explain the apparent epidemics of dengue. Although seasonal factors may have a role, as the literature suggests (10,11), seasonality or other temporal drivers such as fluctuation in the intensity of the country's vector control program are not necessary to explain the qualitative and quantitative patterns of dengue in this equatorial city-state. As our results suggest, the possibility that dengue outbreaks occur in aseasonal locations because of chance should be considered.

This study was funded by the Center for Infectious Disease Epidemiology and Research in the Saw Swee Hock School of Public Health, and additional funding was provided by Singapore's Health Services Research grant number HSRG-0040-2013.

Data used in this paper are available at <http://www.moh.gov.sg>.

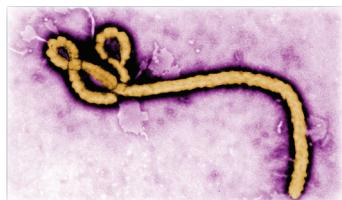
Ms. Chen is a research assistant and doctoral student at the National University of Singapore. Her main research interest is modelling of endemic diseases such as dengue and hand, foot and mouth disease.

References

- Cummings DA, Irizarry RA, Huang NE, Endy TP, Nisalak A, Ungchusak K, et al. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. *Nature*. 2004;427:344–7. <http://dx.doi.org/10.1038/nature02225>
- Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. *Nature*. 2013;496:504–7. <http://dx.doi.org/10.1038/nature12060>
- World Health Organization. Dengue and severe dengue. 2014 Mar [cited 2015 Apr 1]. <http://www.who.int/mediacentre/factsheets/fs117/en/index.html>
- Ooi EE, Goh KT, Gubler DJ. Dengue prevention and 35 years of vector control in Singapore. *Emerg Infect Dis*. 2006;12:887–93. <http://dx.doi.org/10.3201/eid1206.051210>
- Wongkoon S, Jaroensutasinee M, Jaroensutasinee K. Distribution, seasonal variation & dengue transmission prediction in Sisaket, Thailand. *Indian J Med Res*. 2013;138:347–53.
- Bhatia R, Dash AP, Sunyoto T. Changing epidemiology of dengue in South-East Asia. *World Health Organ South East Asia J Public Health*. 2013;2:23–7.
- Ministry of Health Singapore. Communicable diseases surveillance in Singapore 2011. 2012 [updated 2013 Jul 26] [cited 2014 Jan 17]. http://www.moh.gov.sg/content/moh_web/home/Publications/Reports/2012/_communicable_diseasesurveillanceinsingapore2011.html
- Ministry of Health Singapore. Communicable diseases surveillance in Singapore 2012. 2013 [updated 2014 Jan 22] [cited 2014 Jan 17]. http://www.moh.gov.sg/content/moh_web/home/Publications/Reports/2013/Communicable_Diseases_Surveillance_in_Singapore_2012.html
- Gubler DJ. Dengue, urbanization and globalization: the unholy trinity of the 21(st) century. *Trop Med Health*. 2011;39 (Suppl):S3–11. <http://dx.doi.org/10.2149/tmh.2011-S05>
- Hii YL, Rocklöv J, Nawi N, Tang CS, Pang FY, Sauerborn R. Climate variability and increase in intensity and magnitude of dengue incidence in Singapore. *Glob Health Action*. 2009;2.
- Pinto E, Coelho M, Oliver L, Massad E. The influence of climate variables on dengue in Singapore. *Int J Environ Health Res*. 2011;21:415–26. <http://dx.doi.org/10.1080/09603123.2011.572279>
- Horta MA, Bruniera R, Ker F, Catita C, Ferreira AP. Temporal relationship between environmental factors and the occurrence of dengue fever. *Int J Environ Health Res*. 2014;24:471–81. <http://dx.doi.org/10.1080/09603123.2013.865713>
- Fung K. Numbers rule your world: the hidden influence of probabilities and statistics on everything you do. New York: McGraw Hill; 2010.
- Blastland M, Dilnot A. The numbers game: the commonsense guide to understanding numbers in the news, in politics, and in life. New York: Gotham; 2010.
- Gilovich T, Vallone R, Tversky A. The hot hand in basketball: on the misperception of random sequences. *Cogn Psychol*. 1985;17:295–314. [http://dx.doi.org/10.1016/0010-0285\(85\)90010-6](http://dx.doi.org/10.1016/0010-0285(85)90010-6)

Address for correspondence: Alex R. Cook, Saw Swee Hock School of Public Health, Tahir Foundation Building, National University of Singapore, 12 Science Drive 2, Singapore 117549; email: alex.richard.cook@gmail.com

Biomarker Correlates of Survival in Pediatric Patients with Ebola Virus Disease



Dr. Mike Miller reads an abridged version of the article, **Biomarker Correlates of Survival in Pediatric Patients with Ebola Virus Disease.**



<http://www2c.cdc.gov/podcasts/player.asp?f=8633631>

Randomness and Dengue Outbreaks on the Equator

Technical Appendix

Statistical Modeling Approach

The aseasonal model described in the paper belongs to a group of regression models called *autoregressive models*. In these models, the number of dengue cases during 1 week is regressed upon the incidence in preceding weeks. The order of the model (i.e., the number of preceding weeks to autoregress upon) is selected to ensure that there are no further correlations between the model “errors,” or differences between observed and predicted numbers of cases. This process of selecting the order of the model was done by using the Portmanteau, or Ljung-Box, statistical test (I), which suggested the number of dengue cases needed to depend on the weekly number of cases during the previous 4 weeks to meet this statistical requirement. (Using a simpler model that depended on just the previous week did not substantively change the findings in the paper.) To account for incidence necessarily being positive, we used the standard approach of logging the number of dengue cases before modeling them; this approach had a secondary benefit of making the variability more stable. Letting the number of cases in week t be D_t , we therefore calculated and worked with $y_t = \ln(D_t)$. The fitted model was

$$y_t = 0.29 + 0.62y_{t-1} + 0.32y_{t-2} + 0.13y_{t-3} - 0.14y_{t-4} + e_t$$

where e_t is random noise assumed to follow a normal distribution with mean 0 and SD 0.22 (the latter estimated from the data). These estimates were obtained by using the standard regression method of least-squares. The model therefore captures the decay in the risk of secondary cases over the 4 weeks after infection.

The fitted model was used to simulate synthetic datasets that, because the model contains no seasonal forcing, are governed purely by randomness (the e_t terms) and short-term contagion (the y_{t-1} to y_{t-4} terms) but by no other drivers. Three examples are provided in Figure 1 in the main text. From these simulated data, which covered thousands of hypothetical decades of

dengue in Singapore, summaries were extracted that could be compared to the analogous summary from reality. These summaries provide a way to falsify the aseasonality model if the actual summary falls outside the range from the simulations. The summaries were inspired by metrics used in “the hot hand in basketball” perception of randomness study (2) but adapted to the context of weekly case counts of dengue. These metrics included the following:

1. **The autocorrelation function.** This function is the correlation between the (log transformed) number of cases for 1 week against the number k weeks later (i.e., between y_t and y_{t+k}). This function drops from 1 for small lags k to 0 for large lags. The speed at which it drops indicates the degree of memory in the dengue time series (e.g., if there is a high incidence of dengue 1 week, there should be a high incidence during the following week). Each simulation yielded a single curve, over which the curve for the data was overlaid.
2. **The cumulative probabilities.** For a particular number of cases, say D , this metric is the proportion of weeks with D or fewer cases. It therefore captures the overall distribution of dengue incidence but more smoothly than a histogram would. Again, each simulation yielded a single curve, over which the curve for the data was overlaid.
3. **The conditional probability of a rise** in the number of dengue cases from 1 week to the next (t to $t + 1$), given that a rise occurred between week $t - 1$ and week t . Similar probabilities given 2 or 3 successive rises, or 1 to 3 falls, were also considered. These probabilities were calculated directly from the simulations, summarized with a 95% CI, and the corresponding proportions in the historical time series overlaid.
4. **The yearly number of cases.** The distribution from simulations was summarized by using the statistical technique called *kernel density estimation*, and the 10 yearly counts from the data were overlaid.
5. **The maximum weekly number of cases over 10 years.** Again, simulated maxima were summarized by using kernel density estimation, with the maximum weekly incidence observed over the decade in question overlaid as a point.

We also created a further statistical test that counted the number of *runs* (i.e., the number of weeks of successive rises or falls in dengue incidence). The distribution of the number of runs was obtained from the set of simulated datasets and used to generate a p value for the hypothesis that the aseasonal model generated the observed dataset.

The rationale for these summaries is that they capture more complex properties that emerge from the simple aseasonal model described above but that should not concord with the data if seasonal drivers were needed to explain the patterns observed. For instance, with clear seasonal epidemics, following a run of week-on-week rises in incidence, a further rise would be noticeably more likely than 50%, whereas a random process (like the aseasonal model) would see this probability much closer to 50%.

In subtropical Taiwan, for example, there is clear seasonality, which is reflected in cycles in dengue epidemics (Technical Appendix Figure 1, panel A). Simulations from the aseasonal model (Technical Appendix Figure 1, panels B–D) fail to capture both the seasonality and size of dengue outbreaks. The 5 measures we developed were able to indicate clear discrepancies between the data and the simulations (Technical Appendix Figure 2, panels A and C). For equatorial Singapore, however, for all the characteristics considered, the historic data were consistent with every metric we devised to test the aseasonal model:

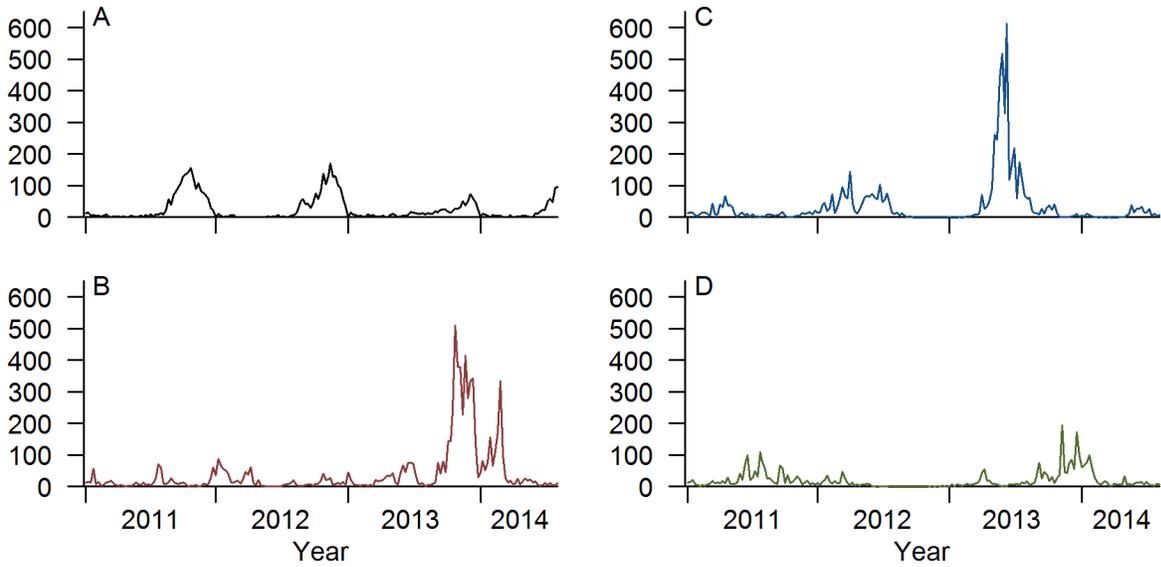
1. For the autocorrelation function (Technical Appendix Figure 2, panel A), the actual autocorrelation falls near the middle of the distribution of simulations under the aseasonal model for all weeks considered, up to the point when the autocorrelation reached 0 (around 6 months). The close correspondence between simulated and observed autocorrelation functions indicates that the random, aseasonal model exhibits the same degree of *memory* as the actual time series; that is, following an epidemic peak, the outbreak reverts to endemic levels at the correct pace; from a trough, epidemics occur at the correct speed.
2. Similarly, the observed cumulative probability function (Technical Appendix Figure 2, panel B) falls near the middle of the distribution of model predictions for all values of the weekly incidence up to the maximum weekly incidence observed. This function indicates that the aseasonal model gives the correct frequency and size of endemic and epidemic phases.

3. The conditional probabilities of rises in the weekly number of cases from 1 week to the next given runs of rises or falls (Technical Appendix Figure 2, panel C) was also consistent between data and model and in the opposite direction to what would be expected given seasonal forcing, with a fall more likely to follow a rise and vice versa. With seasonal forcing, we would expect positive correlations in weather conditions from 1 week to the next to translate to positive correlations in epidemic growth or decline.
4. The 10 annual incidences (Technical Appendix Figure 2, panel D) were fully consistent with the distribution from the model predictions. These incidences indicate that the aseasonal model gives annual tallies that are consistent with those we observed.
5. The peak number of weekly cases over 10 years (Technical Appendix Figure 2, panel E) is similarly within the range of plausible scenarios under the aseasonal model.

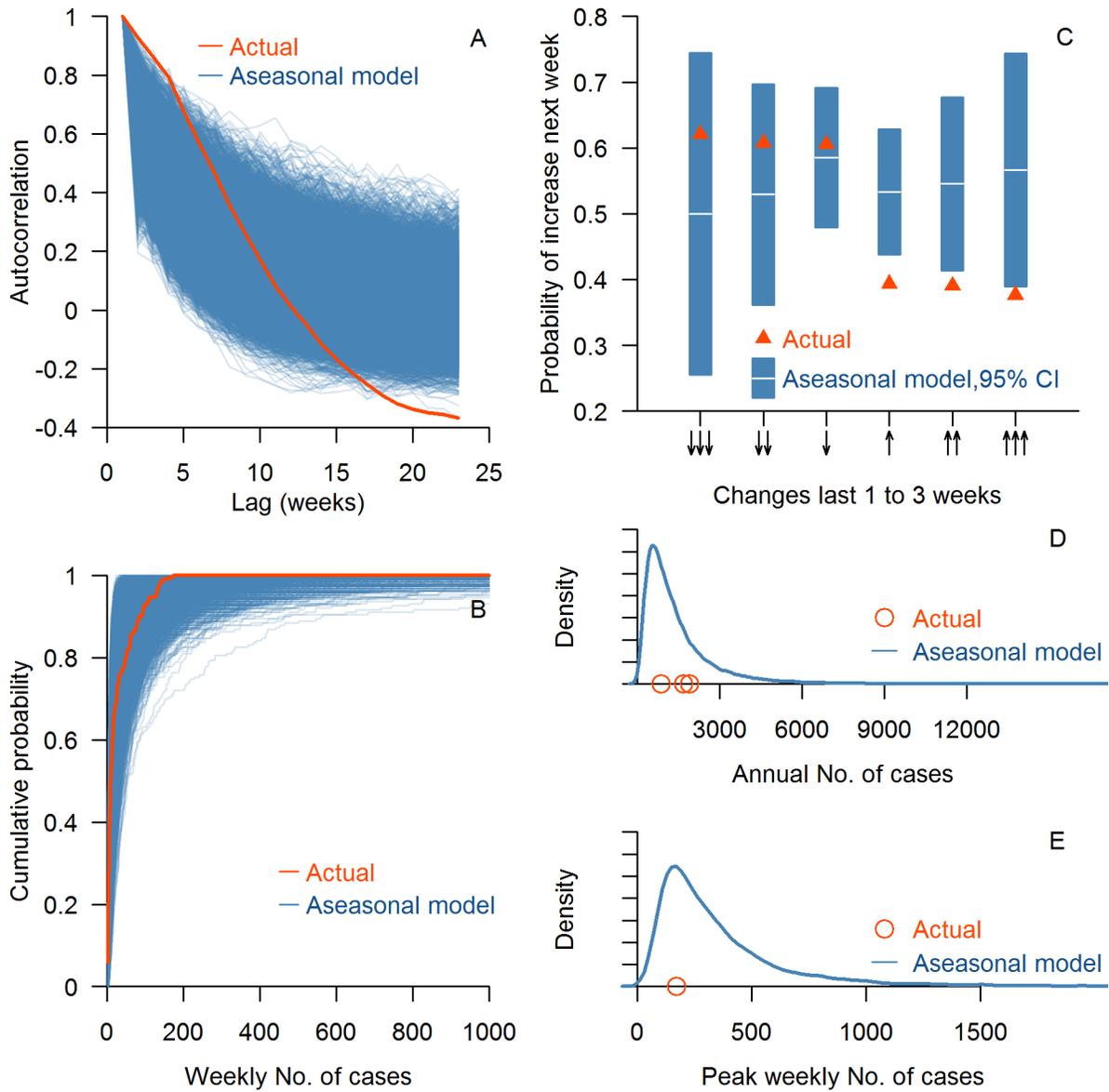
The test for the number of runs did not find any deviation from the aseasonal model ($p = 0.18$), indicating that the consistency of epidemic rise and fall from 1 week to the next also agrees with the absence of seasonality. Further, simulated epidemics (Technical Appendix Figure 1) expressed qualitatively similar behavior to the historic patterns, with similar magnitude and frequency of epidemics and similar endemic behavior.

References

1. Ljung G, Box G. On a measure of lack of fit in time series models. *Biometrika*. 1978;65:297–303.</jrn>
2. Gilovich T, Vallone R, Tversky A. The hot hand in basketball: on the misperception of random sequences. *Cognit Psychol*. 1985;17:295–314.</jrn>



Technical Appendix Figure 1. The aseasonal model applied to Taiwan (2011–2014). A) Dengue trend over >3 years in Taiwan. Clear seasonality can be observed, and epidemics tend to appear at year end. B–D) Aseasonal simulation models. The timing and size of outbreaks differ markedly from the actual scenario of observed dengue incidence.



Technical Appendix Figure 2. Comparison of observed dengue incidence and the simulated aseasonal model in Taiwan. A) Distribution of simulated autocorrelation functions and actual autocorrelation functions at different lags. B) Distribution of cumulative density function of the simulated weekly number of cases and cumulative density function of the actual numbers. C) Conditional probabilities of an increase in number of dengue cases and 95% CIs for simulated scenario and actual data, given 1–3 consecutive decreases or increases. D) Density plot of simulated yearly number of cases and actual number. E) Density plot of simulated 10-year maximum and actual number.