Title: Empirical dynamic modeling reveals ecological drivers of dengue dynamics

Authors: Nicole Nova[1*], Ethan R. Deyle[2,3], Marta S. Shocket[1,4], Andrew J. MacDonald[1,5], Marissa L. Childs[6], Martin Rypdal[7], George Sugihara[2], and Erin A. Mordecai[1]

Affiliations: [1]Department of Biology, Stanford University, Stanford, CA, USA.

[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA.

[3]Department of Biology, Boston University, Boston, MA, USA.

[4]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA.

[5]Earth Research Institute & Bren School of Environmental Science and Management, University of California Santa Barbara, Santa Barbara, CA, USA.

[6]Emmett Interdisciplinary Program in Environment and Resources, Stanford University, Stanford, CA, USA.

[7]Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway.

[*]Correspondence email: nicole.nova@stanford.edu

Keywords: Arbovirus, causality, climate, dengue, empirical dynamic modeling, mosquito-borne disease, rainfall, susceptible population size, temperature.

Abstract

Understanding ecological drivers of mosquito-borne disease is an ongoing public health priority. Previous experiments and mechanistic models suggest mosquito-borne disease transmission involves complex nonlinear interactions between climate and population dynamics. This makes detecting environmental disease drivers at the population level challenging. By analyzing incidence data, estimated susceptible population size, and climate data with methods based on nonlinear time series analysis, collectively referred to as empirical dynamic modeling (EDM), we identified drivers and their interactive effects on dengue dynamics in San Juan, Puerto Rico. Estimated susceptible population size was the strongest causal driver of dengue incidence, and climatic forcing became important above a certain susceptible population size (temperature and rainfall having net positive and negative effects, respectively). Our EDM protocol for measuring and predicting how climate and population dynamics interact to drive epidemics adds to a growing body of empirical studies of complex, nonlinear systems embedded in changing environments.

INTRODUCTION

Mosquito-borne diseases, and dengue in particular, are (re)emerging globally and spreading to higher latitudes in concert with globalization and climate change (Kilpatrick & Randolph 2012; Ryan *et al.* 2019). Dengue virus—transmitted primarily by urban *Aedes aegypti* (Kraemer *et al.* 2015b)—places half of the global human population in 128 countries at infection risk (Brady *et al.* 2012; Kraemer *et al.* 2019) and causes an estimated 390 million annual cases of dengue fever worldwide (Bhatt *et al.* 2013). In the absence of effective vaccines or treatments (Katzelnick *et al.* 2017b, a; Sridhar *et al.* 2018), public health agencies rely on vector control to reduce dengue transmission (Erlanger *et al.* 2008). Effective vector control measures require understanding the mechanisms linking vector ecology and epidemics to better predict disease outbreaks—a major research challenge. Previous prediction models used phenomenological and mechanistic equation-based approaches (Johansson *et al.* 2009; Hii *et al.* 2012; Tran *et al.* 2013; Liu-Helmersson *et al.* 2014; Morin *et al.* 2015; Mordecai *et al.* 2017; Johnson *et al.* 2018), which may not fully capture interdependence between predictors. Here, we used a mechanistic, equation-free, data-driven approach that accounts for interdependence to determine ecological drivers and predict dengue outbreaks in an urban environment.

Since mosquitoes have been shown to be sensitive to climate variables, such as temperature and rainfall (Ibarra *et al.* 2013; Mordecai *et al.* 2019), it is reasonable to ask whether temperature and rainfall are important drivers of dengue outbreaks.

59 Although temperature directly affects mosquito and viral traits in laboratory
60 experiments (Watts *et al.* 1987; Lambrechts *et al.* 2011), the relationship between
61 temperature and dengue incidence in the field has been ambiguous. Thus,
62 temperature-dependent models have had mixed success predicting the timing and
63 magnitudes of epidemics (Hii *et al.* 2012; Johansson *et al.* 2016; Johnson *et al.* 2018).
64 Similarly, the rainfall–dengue relationship is complex, because the effect of rainfall
65 on mosquitoes depends on local breeding habitat and human behavior. In some
66 settings, rainfall provides more container breeding habitat for mosquitoes, thus
67 increasing mosquito abundance and dengue incidence (Ibarra *et al.* 2013). By
68 contrast, low rainfall levels could also facilitate dengue transmission by promoting
69 water storage that serves as standing-water habitat for mosquito breeding
70 (Oliveira-lima *et al.* 2000). Further, heavy rainfall can reduce mosquito abundance
71 by flushing out larvae (Koenraadt & Harrington 2008). The net effect of climate on
72 dengue is a convolution of many different mechanisms of action, and the net
73 outcome depends on specific context details. Moreover, as these details change, the
74 climate–dengue relationship can change.

75 Previous studies using statistical (Johansson *et al.* 2009; Hii *et al.* 2012; Liu-
76 Helmersson *et al.* 2014; Johnson *et al.* 2018) and mechanistic models (Tran *et al.*
77 2013; Morin *et al.* 2015; Mordecai *et al.* 2017) suggest that temperature and rainfall
78 drive dengue transmission in the field. However, these models rely on parameter
79 estimates from laboratory studies that are engineered to isolate single mechanisms
80 of action, hence producing separable relationships between drivers and dengue,
81 eliminating the complex interdependence at the population level. While laboratory

82 studies provide robust validation of a mechanism (Lambrechts *et al.* 2011), the fixed
83 relationships (correlations) taken from them do not necessarily translate into
84 proper causal inference for nonlinear field systems (Sugihara *et al.* 2012). Even if
85 causality exists between two variables in such a system, their correlation can switch
86 signs during different time periods, resulting in a net correlation of zero (Deyle *et al.*
87 2016b). This temporal variation in the direction of correlation results from the
88 nonlinear, state-dependent relationship between the variables (i.e., the importance
89 or direction of one effect depends on the state of another variable). Conversely, even
90 if two variables are consistently correlated, the association could be spurious due to
91 a confounder—a third variable that drives two otherwise unrelated variables. Thus,
92 covariation among variables poses a problem for identifying causal drivers. Both
93 temperature and rainfall follow seasonal patterns in most regions of the world and
94 often covary, making it difficult to separate their effects. Since ecological systems
95 are often nonlinear with covarying drivers, it is difficult to isolate causality from
96 field data.

97 Disease incidence also depends nonlinearly on (potentially climate-driven)
98 transmission rates, because epidemic growth slows as the population of susceptible
99 individuals is exhausted (Rypdal & Sugihara 2019). Thus, susceptible population
100 size is an important driver of infectious disease dynamics (Anderson & May 1979;
101 Kraemer *et al.* 2015a; Rypdal & Sugihara 2019). Variation in susceptible host
102 availability may influence the effects of climate on dengue dynamics, but such
103 interactive effects are difficult to detect in observational data. A lack of rigorous
104 methods has hindered research on potential interdependence of climate and

susceptible population size as drivers of dengue incidence. Because correlations are unreliable and variables are interdependent in nature, disentangling climate from other drivers in disease systems remains challenging (Thai & Anders 2011; Morin *et al.* 2013).

To overcome these challenges, we used empirical dynamic modeling (EDM) (Sugihara *et al.* 2012) to identify and model causal mechanisms driving dengue epidemics (see http://tinyurl.com/EDM-intro for an introduction). EDM differs from statistical and equation-based mechanistic models in two key ways. First, unlike statistical approaches where inferences about cause-and-effect relationships are based on fixed independent pairwise associations between system variables, EDM is based on reconstructing the system *dynamics*. This allows relationships among system variables to change through time to reflect that interactions among variables are changing. Second, unlike equation-based mechanistic models, EDM does not require assumptions about the functional form of the model, but instead reconstructs the dynamic attractor empirically from time series observations. An attractor is a geometric object (i.e., curve or manifold) that embodies the rules for how relationships among variables change with respect to each other through time depending on system state (specific location on the attractor). Like a set of equations, the geometric attractor encompasses the dynamics of a system, and thus can provide an agnostic (without an assumed set of equations) yet mechanistic understanding of the system that is empirically based. Although traditional equation-based mechanistic models can be constructed to account for nonlinearity, these approaches require *a priori* assumptions about the identity and the form of

the causal relationships between variables that may not hold; however, EDM does not rely on such assumptions.

Here, we use EDM and a recent approach for inferring the susceptible population size from incidence data (Rypdal & Sugihara 2019) to answer three questions: (1) Do temperature, rainfall and/or inferred susceptible population size drive dengue incidence? (2) Can we predict dengue dynamics using climate data and inferred susceptible population size? (3) What is the approximate pattern of each causal relationship, and how is this relationship influenced by the other drivers? We examine the hypothesis that temperature and rainfall drive dengue incidence in complex ways that depend on susceptible population size. The EDM protocol presented here—identifying drivers, their predictive power, and state-dependent functional responses—can be applied to other complex, ecological systems.

METHODS

Data and proxy for susceptible population size

We obtained time series of weekly observations of dengue incidence (total number of new dengue cases of all serotypes), temperature (°C), and rainfall (mm) in San Juan, Puerto Rico, 1990–2009 (Figure 1a–c) from the National Oceanic and Atmospheric Administration on November 10, 2016 (http://dengueforecasting.noaa.gov/). Direct measurement of susceptible populations is not feasible, so we used a recently developed method for inferring a proxy for the susceptible population size (Rypdal & Sugihara 2019). This method uses incidence data during the inter-outbreak period to construct a time series for the susceptible population. Although few dengue cases occur during the inter-outbreak period, incidence during this time contains information about the susceptible population size in the next outbreak. Because during an inter-outbreak period the disease system fluctuates around a disease-free equilibrium, a linear approximation of the incidence rate can be made where the coefficient, i.e., the leading eigenvalue, λ, scales linearly with the susceptible population size. If we assume a simple Susceptible-Infected-Recovered (SIR) model (Kermack & McKendrick 1927) for the disease system, then

$$\lambda = \beta\tilde{S} - \gamma \tag{1}$$

where $\tilde{S}$ is a parameter representing the susceptible population size, β is the force of infection, and γ is the recovery rate. Thus, λ is linearly related to the proportion of

the population that is susceptible over time, and can be used as a proxy (see Supporting Information for details).

Since the weekly incidence data are discrete we need to obtain the discrete-time eigenvalue $\tilde{\lambda} = e^{\lambda \Delta t}$. To infer $\tilde{\lambda}$ from the weekly incidence data $I(t)$, we performed linear regression by fitting the statistical model $I(t + \Delta t) = \tilde{\lambda}\, I(t)$ for 12 time points in a 12-week running window ($\Delta t = 1$ week). The model is robust to the window size (Rypdal & Sugihara 2019). In the discrete case, when $\tilde{\lambda} < 1$ the system is stable (inter-outbreak period) and when $\tilde{\lambda} \geq 1$ then the system is unstable (outbreak period). Here, we treated the resulting time series of $\tilde{\lambda}$ as a proxy for the susceptible population size, or "susceptibles index" for short (Figure 1d).

Statistical analyses

All analyses were conducted in R version 3.5.1 (R Development Core Team 2018). We performed pairwise cross-correlations on the time series to investigate time-lagged relationships between potential drivers (i.e., temperature, rainfall, and susceptibles index) and dengue incidence using the `tseries` package version 0.10-45 (Trapletti & Hornik 2018). We calculated the interannual mean to obtain the seasonal variability for each variable. Determining whether the variables follow seasonal trends is important for EDM analyses. Before performing EDM, we normalized each time series to zero mean and unit variance to remove measurement unit bias, ensuring the variables would be comparable and the attractor would not be distorted.

Empirical dynamic modeling (EDM)

All EDM analyses were performed using package `rEDM` version 0.7.1 (Ye *et al.* 2018). EDM includes approaches to infer a system's mechanistic underpinnings and predict its dynamics. EDM uses time series data of one or more variables to construct an attractor in state space (Figure S1). This is called univariable or multivariable state space reconstruction (SSR) for an attractor built using a single or multiple variables, respectively. Properties of the attractor are assessed to examine characteristics of the system (Deyle & Sugihara 2011).

EDM is sensitive to stochasticity, and should be applied in systems where there is evidence of deterministic dynamics (Cummins *et al.* 2015). Some stochasticity is allowed (e.g., there can be stochastic drivers in the system), but the system cannot be entirely stochastic (i.e., there must be low-dimensional deterministic structure to most variables). To test for low-dimensional deterministic dynamics we performed univariable SSR for each variable and used *simplex projection* (Sugihara & May 1990)—a type of nearest neighbor regression—to check that the prediction skill decreases with time to prediction, an indicator of deterministic dynamics (Figures S2a and S4; see Supporting Information for details). To test for nonlinear state dependence of a variable—the motivation behind EDM—we used the *S-map* test for nonlinearity (Sugihara 1994) (Figures S2b,c and S5; see Supporting Information for details).

201 **EDM: Convergent cross-mapping**

202 We used an EDM approach called convergent cross-mapping (CCM) (Sugihara *et al.*
203 2012) to identify drivers of dengue incidence. If two variables are causally related,
204 then a multivariable attractor—where each variable in the system represents a
205 dimension that traces the dynamics of the system—can be semi-reconstructed using
206 lagged versions of just one of the variables (Figure S1). Based on Takens' Theorem,
207 this univariable "shadow attractor" preserves the structural and dynamic properties
208 of the original multivariable attractor (Takens 1981) [see video
209 https://youtu.be/QQwtrWBwxQg (Sugihara *et al.* 2012)]. The concept behind CCM
210 is that if temperature causes dengue incidence, then information about temperature
211 will be embedded in the dynamics of dengue, such that the shadow attractor
212 produced using only dengue dynamics allows us to accurately reconstruct
213 temperature in the past. However, the converse scenario would not be true: since
214 dengue does not cause temperature, the shadow attractor constructed using
215 temperature data should not contain information to accurately reconstruct dengue
216 incidence (see Supporting Information for details).

217 The critical criterion for testing the existence of causality using CCM is checking that
218 the cross-mapping skill monotonically increases and plateaus (i.e., converges) with
219 the length of the response variable data series used in cross-mapping. The cross-
220 mapping skill, ρ, is the Pearson's correlation coefficient between predicted driver
221 values using the univariable SSR of the response variable, and the observed driver
222 values. We used the Kendall's τ test as a significance test for convergence using the

223 `Kendall` package version 2.2 (McLeod 2011). This test checks whether cross-
224 mapping skill has a significant monotonic increasing trend. If $\tau > 0$ then there is
225 convergence (Grziwotz *et al.* 2018).

226 Based on the cross-correlation analyses (Figure S6), we applied a 9-week time lag
227 between temperature and incidence, an averaged lag of 3–9 weeks for rainfall, and a
228 5-week lag for the susceptibles index. The rainfall variable is a proxy for standing
229 water as mosquito breeding habitat, and thus we were interested in the
230 accumulation of water over time. We refer to these as "ecological lags" and treat
231 them as proxies for the time delay of cause-and-effect, since they account for the
232 ecological processes in the causal chain of events (ecological lags are separate from
233 lags used in SSR). These ecological time lags are consistent with results from other
234 field studies, which showed that temperature and rainfall predict dengue cases 6–12
235 weeks ahead (Chen *et al.* 2010; Ibarra *et al.* 2013).

236 **EDM: Null models**

237 For CCM, we assessed the strength of the evidence for causal effects of potential
238 drivers on dengue using two null models that control for the seasonal trend
239 observed in all variables (Figure 2). These null models address the sensitivity of
240 CCM to periodic fluctuations (i.e., seasonality), which can make two variables appear
241 to be causally linked when instead they are simply synchronized by an unobserved
242 seasonal variable (Cobey & Baskerville 2016; Deyle *et al.* 2016a). In the first null
243 model ("seasonal null model"), we preserved the seasonal signal in the null, but
244 randomized the interannual anomalies, and compared model performance with the

245 true time series (Deyle *et al.* 2016a). In the second, more conservative null model
246 ("Ebisuzaki null model"), we conserved any serial correlation (beyond seasonal) and
247 randomized the phases of Fourier transformed time series, and compared model
248 performance with the true time series (Ebisuzaki 1997).

249 We also repeated the CCM method in the nonsensical, reverse-causal direction (e.g.,
250 to test whether incidence drives climate) as a control for potential spurious
251 relationships generated by non-causal covariation (e.g., due to seasonality). This
252 addresses the issue of synchrony, in which CCM can indicate bidirectional causality
253 when one direction is false or nonsensical (Baskerville & Cobey 2017; Sugihara *et al.*
254 2017).

255 **EDM: Forecast improvement**

256 We examined the predictive power of each driver—or combination of drivers—on
257 dengue incidence by assessing how well we can predict dengue dynamics using
258 temperature, rainfall, and susceptibles index. We used a combination of univariable
259 SSR (i.e., with incidence data) and multivariable SSR to determine the improvement
260 of forecasting (using simplex projection) when including each driver or combination
261 of drivers (Dixon *et al.* 1999; Deyle *et al.* 2013, 2016a).

262 We investigated the potential forecast improvement of dengue incidence using
263 temperature, rainfall, susceptibles index, and their combined effect. In addition, we
264 investigated the predictive power of dengue incidence using multivariable SSR with

just the potential driver variables: temperature, rainfall, and susceptibles index (see Supporting Information).

EDM: Scenario exploration

In nonlinear systems, drivers generally do not have a constant effect. The effect is state-dependent—the strength and possibly the direction of the effect depends on the current state of the system. Scenario exploration with multivariable EDM allowed us to assess the effect of a small change in temperature or rainfall on dengue incidence, across different states of the system. The outcome of these small changes allowed us to deduce the relationship between each climate driver and dengue incidence and how they depend on the system state. For each time step t we used S-maps (Sugihara 1994; Deyle *et al.* 2016a) to predict dengue incidence using a small increase ($+\Delta X/2$) and a small decrease ($-\Delta X/2$) of the observed value of driver $X(t)$ (temperature or rainfall). For each putative climate driver, the difference in dengue predictions between these small changes is $\Delta Y = Y(t+1)\left[X(t) + \frac{\Delta X(t)}{2}\right] - Y(t+1)\left[X(t) - \frac{\Delta X(t)}{2}\right]$, where $Y(t+1)$ is a function of X and all other state variables, and we used $\Delta Y/\Delta X$ to approximate the effect of driver X at time t. We repeated this over all time steps in our time series data for both temperature and rainfall to recover their approximate relationships with dengue incidence at different states of the system. Scenario exploration analyses were repeated across several model parameterizations to address potential sensitivity to parameter settings (see Supporting Information).

286 **RESULTS**

287 **Detecting drivers**

288 EDM showed that temperature, rainfall, and susceptibles index drive dengue
289 incidence since the convergence criterion was met (Kendall's $\tau > 0$, $p < 0.01$) in all
290 three CCM cases (Figure 3a–c). Rainfall and susceptibles index were significant
291 drivers of dengue incidence beyond seasonality, as their effects were
292 distinguishable from seasonal and Ebisuzaki null models (Figures 3b, c and S8b, c; p
293 < 0.05). This implies significant causal effects of both rainfall and susceptible
294 population size on dengue, which are not obscured by a confounder with a periodic
295 signal. However, temperature was not a significant driver beyond seasonality
296 (Figures 3a and S8a; $p > 0.05$). This implies we cannot rule out the possibility that
297 the apparent forcing of temperature on dengue is due to a confounder with a
298 seasonal signal. However, if there is no such confounder, then the seasonal trend in
299 temperature, which accounts for most temperature variation in San Juan, drives the
300 seasonal trend observed in dengue incidence (i.e., seasonality of temperature drives
301 seasonality of dengue). Compared to the other drivers, the predictive skill of the
302 temperature null model was relatively high (Figure 3a), suggesting that temperature
303 seasonality in the null model was predictive. This further supports the notion that
304 seasonal temperature may be driving dengue dynamics.

305 As expected, EDM tests for causality in the nonsensical directions—incidence
306 driving temperature or rainfall—were not significant (i.e., no convergence; Figure

307 S7, black lines). This further supports that temperature and rainfall drive dengue
308 incidence, because their causal relationships were not confounded by spurious
309 bidirectionality. Further, the null models for the nonsensical directions of causality
310 (Figure S7, grey lines) also displayed no convergence (completely flat), as expected
311 (i.e., seasonality of dengue incidence does not drive seasonality of temperature or
312 rainfall). However, seasonality (or any periodicity) of temperature, rainfall and
313 susceptibles index drive dengue dynamics, shown by convergence of the seasonal
314 and Ebisuzaki null models (grey lines in Figures 3 and S8).

315 **Predictive power of drivers**

316 Dengue incidence was highly predictable using univariable SSR of incidence data
317 alone (Adjusted R^2 = 0.8922, ρ = 0.9446; Figure 4a). The predictive power of dengue
318 incidence improved only slightly when temperature, rainfall, and susceptibles index
319 were added in a multivariable SSR model (Adjusted R^2 = 0.8927, ρ = 0.9448; Figure
320 4d). These results suggest that dengue incidence time series contain information
321 about these drivers, resulting in the high predictability of the attractor in
322 univariable SSR (Figure 4a).

323 Dengue dynamics were also highly predictable using only the driver time series (i.e.,
324 temperature, rainfall, and susceptibles index) in a multivariable SSR model
325 (Adjusted R^2 = 0.5044, ρ = 0.7102; Figure 4c), where timing and magnitude of
326 epidemics were captured reasonably well. However, the model using only
327 temperature and rainfall data did not predict dengue incidence as well (Adjusted R^2
328 = 0.0533, ρ = 0.2309; Figure 4b). Thus, adding the susceptibles index increases

329 predictive power compared to climate variables alone, particularly for predicting
330 the magnitude of outbreaks (Figure 4c). However, temperature and rainfall did
331 capture the timing (seasonality) of the epidemics (Figure 4b). This supports the
332 notion that seasonality of temperature and rainfall is important for explaining the
333 seasonality of dengue, and that susceptible population size is important for
334 determining epidemic size. All SSR models (Figure 4a–d) had significant ($p < 0.001$)
335 F statistics >> 1 from ANOVA, rejecting the null hypothesis (i.e., no relationship
336 between predicted and observed dengue incidence).

337 **State-dependent functional responses**

338 By investigating the rate of change of dengue incidence as a function of climate
339 across system states, we found that temperature had a small *positive* median effect
340 (0.0035, Wilcox $p < 0.001$) on dengue incidence. A positive effect is expected
341 (Mordecai *et al.* 2017) for the temperature range in Puerto Rico (Figure S9, black
342 dashed lines), although the effect was occasionally much stronger, both positive and
343 negative (Figure 5a, b). The large negative effects occurred only at the highest
344 temperature values (as predicted by mechanistic models of temperature-dependent
345 transmission), reinforced by a lower quantile regression with a strongly negative
346 slope (Figure 5b, bottom dashed red line). However, positive effects occurred across
347 the whole temperature range.

348 Rainfall had a small *negative* median effect (–0.0022, Wilcox $p < 0.001$), but
349 occasionally had very large negative effects (Figure 5a, c). These large, negative
350 effects of rainfall on dengue occurred when there was less than 100 mm of rain per

351 week (Figure 5c), in line with expectations that *low* amounts of rainfall could lead to
352 *high* number of dengue cases, since people tend to store water in containers during
353 periods of drought, enabling mosquitoes to breed (Oliveira-lima *et al.* 2000). There
354 are also small positive effects of rainfall on dengue (Figure 5c), suggesting that
355 overall the results showed competing effects of low–moderate rain providing
356 standing water for mosquito breeding and humans storing standing water where
357 mosquitoes can breed when there is drought or low rain.

358 The results on climate effects suggest that the strength and direction of the effects of
359 climate on dengue dynamics depend on the state of the system. A potential cause of
360 state-dependent climate effects on dengue dynamics is the variation in the
361 susceptible population size over time (Figure 6a, b). Outbreaks do not occur when
362 there are too few susceptible people to get infected. As expected, when the
363 susceptible population size was small ($\tilde{\lambda} < 0.85$) incidence was insensitive to
364 climate (Figure 6c, e). By contrast, when the susceptible population size was large
365 ($\tilde{\lambda} > 0.85$), temperature and rainfall effects on dengue incidence appeared (Figure
366 6d, f). The gradual increase and decrease of the rate of change of dengue as a
367 function of temperature (Figure 6d, red solid lines) aligned well with the sigmoidal
368 part (Figure S9, black dashed lines representing the temperature range in our
369 study) of the unimodal temperature response curve (assuming transmission
370 primarily by *Ae. aegypti*) developed previously (Mordecai *et al.* 2017). This is an
371 important finding, since evidence of climate functional responses for disease
372 dynamics is rare due to the difficulty of obtaining appropriately informative field

data. It is possible that if we had temperature data ranging across a larger spectrum—possibly by assembling data across multiple climates—that the empirical functional response derived from EDM would also look unimodal.

DISCUSSION

EDM can identify causal drivers, quantify their predictive power, and approximate functional responses in nonlinear, state-dependent ecological systems, which few other methods can do. Infectious disease dynamics are often seasonal, adding another layer of complexity. Here, we used seasonal null models and a well-studied ecological system to disentangle nonlinear and interactive mechanisms driving disease dynamics. First, EDM detected rainfall, susceptible population size, and plausibly temperature (mostly via its seasonality) as drivers of dengue incidence. The seasonal variation in incidence was more attributed to climate, while the interannual variation in incidence was more explained by the susceptible population (Figure 3). Second, EDM provided a predictive model based on these three drivers that had a reasonably good fit to dengue incidence data ($R^2 = 0.50$, $\rho = 0.71$; Figure 4c). Dengue dynamics were also highly predictable from incidence data alone ($R^2 = 0.89$, $\rho = 0.94$; Figure 4a); thus, robust to missing state variables. This implies that EDM methods could be powerful for forecasting epidemics, provided that surveillance efforts continue to report weekly case data. Third, EDM revealed that climate effects on dengue appeared once the susceptible population size exceeded a threshold ($\tilde{\lambda} > 0.85$; Figure 6).

The fact that climate effects are observed before the onset of an outbreak, when $\tilde{\lambda} = 1$ (Rypdal & Sugihara 2019), suggests that rainfall, and possibly temperature, have an effect on the timing of an impending epidemic. Thus, although rainfall and temperature might not influence the dynamics of dengue during an inter-outbreak

398 period when the susceptible population size is small, climate could act as a catalyst
399 to spark an epidemic once the susceptible population size is large enough. This
400 resonates with the notion that climate could drive the force of infection, β, thus
401 influencing the susceptibles index, λ (Rypdal & Sugihara 2019). The timing of an
402 outbreak, when $\lambda \geq 0$ (or in the discrete case when $\tilde{\lambda} \geq 1$), could be attributed to
403 the changes in β caused by seasonal climatic drivers (Rypdal & Sugihara 2019).
404 Further, seasonality of temperature and rainfall had higher predictive skill than
405 seasonality of susceptibles index (Figures 3 and S8, grey solid lines); however,
406 adding susceptibles index dramatically improved dengue forecasts due to more
407 accurate epidemic magnitudes (Figure 4b, c). Thus, climate may be mostly
408 responsible for the *timing* of seasonal epidemics, while susceptible population size
409 may mostly determine the epidemic *magnitude*. Using the same dataset, Johnson *et*
410 *al.* (2018) found that mechanistic models could predict the timing of seasonal
411 epidemics, but that a phenomenological machine learning component was needed to
412 capture interannual variation in epidemic magnitude. Our work suggests that the
413 unobserved size of the susceptible population was a key missing link for predicting
414 magnitude variation across years.

415 We showed that susceptible population size is an important modifier of climate
416 effects on dengue (Figure 6). This climate–susceptible population interdependence
417 might be generalizable across other infectious diseases. For example, for influenza
418 dynamics, population density in cities—potentially a proxy for susceptible
419 population density—modulated climate effects on disease transmission (Dalziel *et*
420 *al.* 2018). Given complex immune and serotype dynamics of dengue (Katzelnick *et*

421 *al.* 2017b), total population density might not work as a proxy for the susceptible
422 population density in dengue dynamics. It has been difficult for previous
423 mechanistic models to capture susceptible dynamics for dengue. By inferring the
424 susceptibles index from incidence data (Rypdal & Sugihara 2019), we were able to
425 capture the strong influence of the susceptible population size on dengue dynamics,
426 which in turn moderated the effect of climate drivers on dengue dynamics. We
427 showed that climate–dengue relationships were only detectable once we accounted
428 for seasonal and interannual variation in susceptible availability (Figure 6d, f). This
429 is expected from theory (Kermack & McKendrick 1927; Xu *et al.* 2017), but
430 demonstrating it empirically is a unique contribution of this study.

431 Even when accounting for susceptible availability, the effects of temperature and
432 rainfall on dengue were still strongly state-dependent (Figure 6d, f). The remaining
433 variation in temperature and rainfall effects on dengue—given that the susceptible
434 population is large enough for an outbreak—may be partially explained by variation
435 in temperature and rainfall over time and space that is not captured by weekly
436 climate averages, and by interactions between temperature and rainfall. Further,
437 any subtle lagged effects of temperature or rainfall on dengue are not captured by a
438 single ecological lag (a 9-week or an averaged 3–9-week lag, respectively). We
439 expect multiple time lags to operate based on the different ecological processes that
440 generate reported cases: mosquito population growth, mosquitoes getting infected
441 and biting susceptible humans, development of symptoms, and care-seeking. The
442 ecological lags are simple proxies for this complex set of ecological processes.

443 EDM is like any other quantitative method in that it is appropriately applied when
444 its assumptions are met. EDM is sensitive to stochasticity and synchrony (e.g., via
445 seasonality), so it is important to first determine whether a system is governed by
446 deterministic dynamics, and then to use null models that account for synchrony
447 (Ebisuzaki 1997; Deyle *et al.* 2016a; Chang *et al.* 2017). Nevertheless, EDM is a
448 useful approach for testing causality between hypothesized drivers and response
449 variables in a dynamic system, and recent work allows EDM to perform well with
450 stochasticity (Cenci & Saavedra 2018; Cenci *et al.* 2019). To infer strongly supported
451 causal relationships, it is also important to consider the mechanistic hypotheses
452 underlying the system. For dengue, there are complementary lines of evidence
453 about the drivers of incidence from experimental studies (Watts *et al.* 1987;
454 Lambrechts *et al.* 2011), mechanistic models parameterized with data (Otero *et al.*
455 2006; Mordecai *et al.* 2017), phenomenological studies across settings (Johansson *et*
456 *al.* 2009; Hii *et al.* 2012; Ibarra *et al.* 2013), and EDM from this study. In particular,
457 the 19-year weekly time series dataset used here has been investigated using
458 multiple methods (Johansson *et al.* 2019); however, most of these methods are
459 either not mechanistic or include simple mechanisms that do not account for the
460 joint influence of climate and susceptible dynamics. EDM gives us the opportunity to
461 infer mechanisms, and assess their support in other studies using different
462 approaches. In sum, EDM has limitations, but the limitations of other approaches are
463 at least as great, and may not capture state-dependent effects. When combined with
464 supporting evidence from other approaches, EDM can be a powerful test for

465 causality, prediction, and functional relationships between drivers and response
466 variables.

467 Connecting climate and dengue at the population level is challenging, because causal
468 relationships are likely to be nonlinear and state-dependent. Thus, a toolbox of
469 rigorous methods for testing hypotheses, understanding mechanisms, and making
470 predictions is essential for understanding disease dynamics in complex, natural
471 populations. Ultimately, understanding how climate-driven vector-borne diseases
472 are influenced by other variables, such as susceptible population size, is important
473 for optimizing vector control under critical conditions where climate might spark
474 epidemics. EDM is most appropriate for deterministic systems, which are common
475 in disease ecology. The mechanisms inferred from EDM could be applied to
476 understand and predict future ecological responses to changing environments,
477 including dengue epidemics in a world undergoing global change.

ACKNOWLEDGEMENTS

We thank Giulio De Leo, Marcus Feldman, Dmitri Petrov, and members of the Fukami, Mordecai, Peay, and Sugihara labs for helpful feedback. NN was supported by the Bing Fellowship in Honor of Paul Ehrlich. ERD and GS were supported by the National Science Foundation (NSF) DEB-1655203, NSF-ABI-1667584, DoD-Strategic Environmental Research and Development Program (SERDP) 15 RC-2509; Lenfest Foundation Award 00028335 and the McQuown Chair in Natural Sciences, University of California, San Diego. MSS and EAM were supported by an NSF Ecology and Evolution of Infectious Diseases grant (DEB-1518681). EAM was also supported by an NSF Rapid Response Research grant (RAPID 1640780), the NIH (R35GM133439), the Stanford University Woods Institute for the Environment Environmental Ventures Program, the Hellman Faculty Fellowship, and the Terman Fellowship. AJM was supported by an NSF Postdoctoral Research Fellowship in Biology (1611767). MLC was supported by the Lindsay Family E-IPER Fellowship and Illich-Sadowsky Interdisciplinary Graduate Fellowship.

AUTHORSHIP

NN, EAM, GS, and MR conceived the idea for the study and designed the analyses. MR inferred the proxy for susceptible population size time series. NN, ERD, MSS, AJM, and MLC performed the analyses. NN wrote the first draft of the manuscript. All authors contributed to manuscript revisions and gave approval for publication.

REFERENCES

Anderson, R.M. & May, R.M. (1979). Population biology of infectious diseases: Part I. *Nature*.

Baskerville, E.B. & Cobey, S. (2017). Does influenza drive absolute humidity? *Proc. Natl. Acad. Sci.*, 114, 201700369.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., *et al.* (2013). The global distribution and burden of dengue. *Nature*, 496, 504–507.

Brady, O.J., Gething, P.W., Bhatt, S., Messina, J.P., Brownstein, J.S., Hoen, A.G., *et al.* (2012). Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLoS Negl. Trop. Dis.*, 6, e1760.

Cenci, S. & Saavedra, S. (2018). Uncertainty quantification of the effects of biotic interactions on community dynamics from nonlinear time-series data. *J. R. Soc. Interface*, 15.

Cenci, S., Sugihara, G. & Saavedra, S. (2019). Regularized S-map for inference and forecasting with noisy ecological time series. *Methods Ecol. Evol.*, 2041–210X.13150.

Chang, C.W., Ushio, M. & Hsieh, C. hao. (2017). Empirical dynamic modeling for beginners. *Ecol. Res.*, 32, 785–796.

Chen, S.C., Liao, C.M., Chio, C.P., Chou, H.H., You, S.H. & Cheng, Y.H. (2010). Lagged temperature effect with mosquito transmission potential explains dengue variability in southern Taiwan: Insights from a statistical analysis. *Sci. Total*

Environ., 408, 4069–4075.

Cobey, S. & Baskerville, E.B. (2016). Limits to causal inference with state-space reconstruction for infectious disease. *PLoS One*, 11, 1–22.

Cummins, B., Gedeon, T. & Spendlove, K. (2015). On the Efficacy of State Space Reconstruction Methods in Determining Causality. *SIAM J. Appl. Dyn. Syst.*, 14, 335–381.

Dalziel, B.D., Kissler, S., Gog, J.R., Viboud, C., Bjørnstad, O.N., Metcalf, C.J.E., *et al.* (2018). Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. *Science (80-.).*, 362, 75–79.

Deyle, E.R., Fogarty, M., Hsieh, C., Kaufman, L., Maccall, A.D. & Munch, S.B. (2013). Predicting climate effects on Pacific sardine. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 6430–6435.

Deyle, E.R., Maher, M.C., Hernandez, R.D., Basu, S. & Sugihara, G. (2016a). Global environmental drivers of influenza. *Proc. Natl. Acad. Sci.*, 113, 13081–13086.

Deyle, E.R., May, R.M., Munch, S.B. & Sugihara, G. (2016b). Tracking and forecasting ecosystem interactions in real time. *Proc. R. Soc. B Biol. Sci.*, 283.

Deyle, E.R. & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS One*, 6.

Dixon, P.A., Milicich, M.J. & Sugihara, G. (1999). Episodic fluctuations in larval supply. *Science (80-.).*, 283, 1528–1530.

Ebisuzaki, W. (1997). A method to estimate the statistical significance of a

correlation when the data are serially correlated. *J. Clim.*, 10, 2147–2153.

Erlanger, T.E., Keiser, J. & Utzinger, J. (2008). Effect of dengue vector control interventions on entomological parameters in developing countries: A systematic review and meta-analysis. *Med. Vet. Entomol.*

Grziwotz, F., Strauß, J.F., Hsieh, C. & Telschow, A. (2018). Empirical Dynamic Modelling Identifies different Responses of Aedes Polynesiensis Subpopulations to Natural Environmental Variables. *Sci. Rep.*, 8, 16768.

Hii, Y.L., Zhu, H., Ng, N., Ng, L.C. & Rocklöv, J. (2012). Forecast of Dengue Incidence Using Temperature and Rainfall. *PLoS Negl. Trop. Dis.*, 6, e1908.

Ibarra, A.M.S., Ryan, S.J., Beltrá N, E., Mejía, R.L., Silva, M. & Muñ Oz, N. (2013). Dengue Vector Dynamics (Aedes aegypti) Influenced by Climate and Social Factors in Ecuador: Implications for Targeted Control. *PLoS One*, 8.

Johansson, M.A., Apfeldorf, K.M., Dobson, S., Devita, J., Buczak, A.L., Baugher, B., *et al.* (2019). An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc. Natl. Acad. Sci.*, 116, 201909865.

Johansson, M.A., Dominici, F. & Glass, G.E. (2009). Local and global effects of climate on dengue transmission in Puerto Rico. *PLoS Negl. Trop. Dis.*, 3, e382.

Johansson, M.A., Reich, N.G., Hota, A., Brownstein, J.S. & Santillana, M. (2016). Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Sci. Rep.*, 6, 33707.

Johnson, L.R., Gramacy, R.B., Cohen, J., Mordecai, E., Murdock, C., Rohr, J., *et al.* (2018). Phenomenological forecasting of disease incidence using

562 heteroskedastic gaussian processes: A dengue case study. *Ann. Appl. Stat.*, 12,
563 27–66.

564 Katzelnick, L.C., Coloma, J. & Harris, E. (2017a). Dengue: knowledge gaps, unmet
565 needs, and research priorities. *Lancet Infect. Dis.*, 17, e88–e100.

566 Katzelnick, L.C., Gresh, L., Halloran, M.E., Mercado, J.C., Kuan, G., Gordon, A., *et al.*
567 (2017b). Antibody-dependent enhancement of severe dengue disease in
568 humans. *Science (80-.).*, 358, 929–932.

569 Kermack, W.O. & McKendrick, A.G. (1927). A Contribution to the Mathematical
570 Theory of Epidemics. *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 115, 700–721.

571 Kilpatrick, A.M. & Randolph, S.E. (2012). Drivers, dynamics, and control of emerging
572 vector-borne zoonotic diseases. *Lancet*.

573 Koenraadt, C.J.M. & Harrington, L.C. (2008). Flushing effect of rain on container-
574 inhabiting mosquitoes Aedes aegypti and Culex pipiens (Diptera: Culicidae). *J.*
575 *Med. Entomol.*, 45, 28–35.

576 Kraemer, M.U.G., Perkins, T.A., Cummings, D.A.T., Zakar, R., Hay, S.I., Smith, D.L., *et al.*
577 (2015a). Big city, small world: Density, contact rates, and transmission of
578 dengue across Pakistan. *J. R. Soc. Interface*, 12.

579 Kraemer, M.U.G., Reiner, R.C., Brady, O.J., Messina, J.P., Gilbert, M., Pigott, D.M., *et al.*
580 (2019). Past and future spread of the arbovirus vectors Aedes aegypti and
581 Aedes albopictus. *Nat. Microbiol.*, 4, 854–863.

582 Kraemer, M.U.G., Sinka, M.E., Duda, K.A., Mylne, A.Q.N., Shearer, F.M., Barker, C.M., *et*
583 *al.* (2015b). The global distribution of the arbovirus vectors Aedes aegypti and

584 Ae. Albopictus. *Elife*, 4.

585 Lambrechts, L., Paaijmans, K.P., Fansiri, T., Carrington, L.B., Kramer, L.D., Thomas,
586 M.B., *et al.* (2011). Impact of daily temperature fluctuations on dengue virus
587 transmission by Aedes aegypti. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 1–6.

588 Liu-Helmersson, J., Stenlund, H., Wilder-Smith, A. & Rocklöv, J. (2014). Vectorial
589 capacity of Aedes aegypti: Effects of temperature and implications for global
590 dengue epidemic potential. *PLoS One*, 9.

591 McLeod, A.I. (2011). Kendall: Kendall rank correlation and Mann-Kendall trend test.

592 Mordecai, E.A., Caldwell, J.M., Grossman, M.K., Lippi, C.A., Johnson, L.R., Neira, M., *et*
593 *al.* (2019). Thermal biology of mosquito-borne disease. *Ecol. Lett.*, ele.13335.

594 Mordecai, E.A., Cohen, J.M., Evans, M. V., Gudapati, P., Johnson, L.R., Lippi, C.A., *et al.*
595 (2017). Detecting the impact of temperature on transmission of Zika, dengue,
596 and chikungunya using mechanistic models. *PLoS Negl. Trop. Dis.*, 11,
597 e0005568.

598 Morin, C.W., Comrie, A.C. & Ernst, K. (2013). Climate and dengue transmission:
599 Evidence and implications. *Environ. Health Perspect.*

600 Morin, C.W., Monaghan, A.J., Hayden, M.H., Barrera, R. & Ernst, K. (2015).
601 Meteorologically driven simulations of dengue epidemics in San Juan, PR. *PLoS*
602 *Negl. Trop. Dis.*, 9, e0004002.

603 Oliveira-lima, W., Hodgson, J.C., Pontes, R.J.S. & Freeman, J. (2000). Vector densities
604 that potentiate dengue outbreaks in a Brazilian city. *Trop. Med.*, 62, 378–383.

Otero, M., Solari, H.G. & Schweigmann, N. (2006). A stochastic population dynamics model for Aedes aegypti: Formulation and application to a city with temperate climate. *Bull. Math. Biol.*, 68, 1945–1974.

Ryan, S.J., Carlson, C.J., Mordecai, E.A. & Johnson, L.R. (2019). Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.*, 13, e0007213.

Rypdal, M. & Sugihara, G. (2019). Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics. *Nat. Commun.*, 10, 2374.

Sridhar, S., Luedtke, A., Langevin, E., Zhu, M., Bonaparte, M., Machabert, T., *et al.* (2018). Effect of Dengue Serostatus on Dengue Vaccine Safety and Efficacy. *N. Engl. J. Med.*, 379, 327–340.

Sugihara, G. (1994). Nonlinear Forecasting for the Classification of Natural Time Series. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 348, 477–495.

Sugihara, G., Deyle, E.R. & Ye, H. (2017). Reply to Baskerville and Cobey: Misconceptions about causation with synchrony and seasonal drivers. *Proc. Natl. Acad. Sci.*, 114, 201700998.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., *et al.* (2012). Detecting Causality in Complex Ecosystems. *Science (80-.).*, 338, 496–500.

Sugihara, G. & May, R.M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*.

Takens, F. (1981). Detecting strange attractors in turbulence. In: *Dynamical Systems*

and Turbulence, Warwick 1980, Lecture Notes in Mathematics (eds. Rand, D. & Young, L.-S.). Springer-Verlag, pp. 366–381.

Thai, K.T.D. & Anders, K.L. (2011). The role of climate variability and change in the transmission dynamics and geographic distribution of dengue. *Exp. Biol. Med.*, 236, 944–954.

Tran, A., L'Ambert, G., Lacour, G., Benoît, R., Demarchi, M., Cros, M., *et al.* (2013). A rainfall- and temperature-driven abundance model for Aedes albopictus populations. *Int. J. Environ. Res. Public Health*, 10, 1698–1719.

Trapletti, A. & Hornik, K. (2018). tseries: Time Series Analysis and Computational Finance.

Watts, D.M., Burke, D.S., Harrison, B.A., Whitmire, R.E. & Nisalak, A. (1987). Effect of temperature on the vector efficiency of Aedes aegypti for dengue 2 virus. *Am. J. Trop. Med. Hyg.*, 36, 143–152.

Xu, L., Stige, L.C., Chan, K.-S., Zhou, J., Yang, J., Sang, S., *et al.* (2017). Climate variation drives dengue dynamics. *Proc. Natl. Acad. Sci.*, 114, 113–118.

Ye, H., Clark, A., Deyle, E., Munch, S., Cai, J., Cowles, J., *et al.* (2018). rEDM: Applications of Empirical Dynamic Modeling from Time Series. R package version 0.7.1.

FIGURES

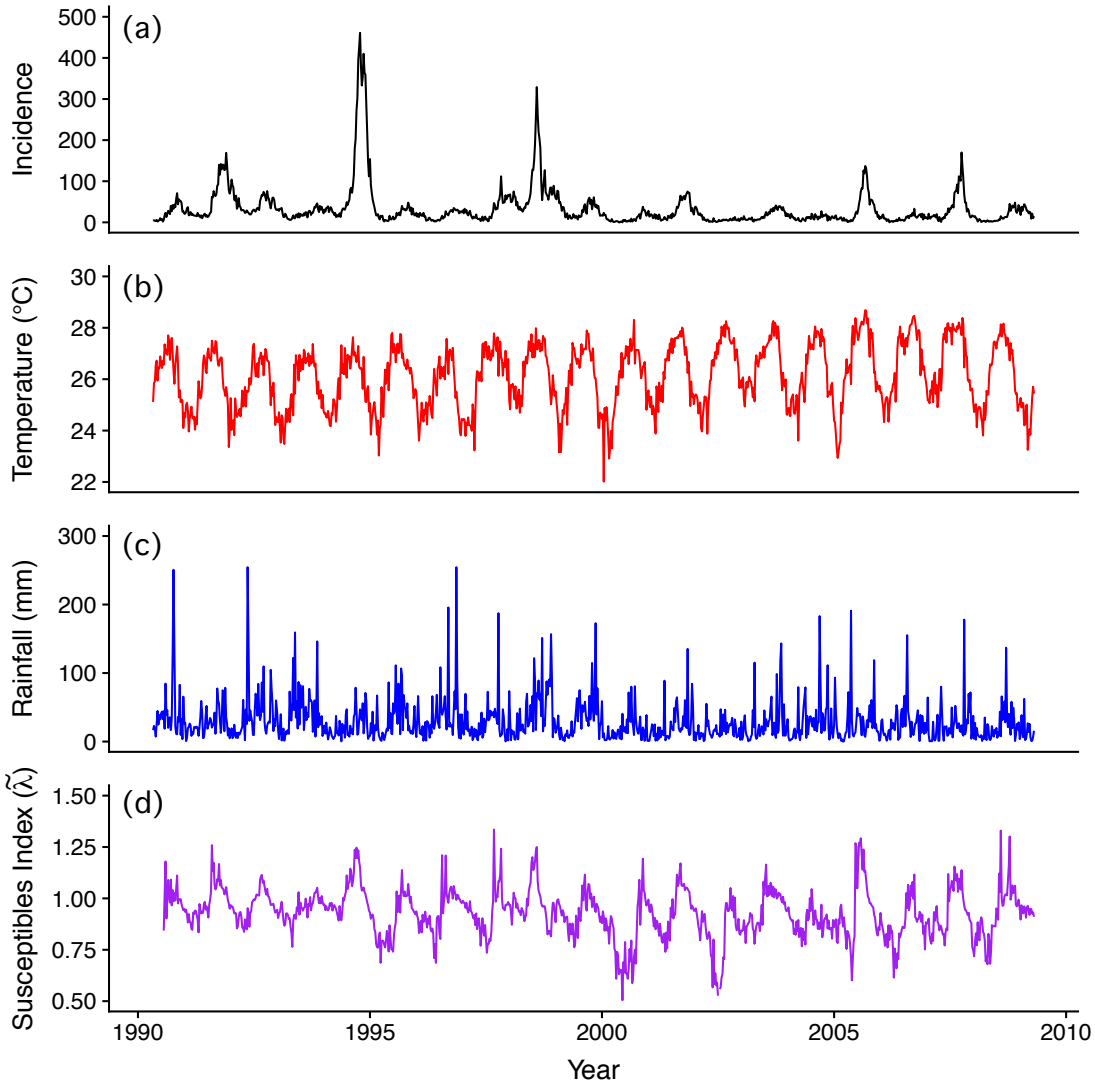


Figure 1. Time series data. Time series (1990–2009) of (a) weekly dengue incidence (i.e., total number of cases per week), (b) weekly average temperature, (c) total weekly rainfall, and (d) a proxy for susceptible population size (see Supporting Information for details) in San Juan, Puerto Rico.

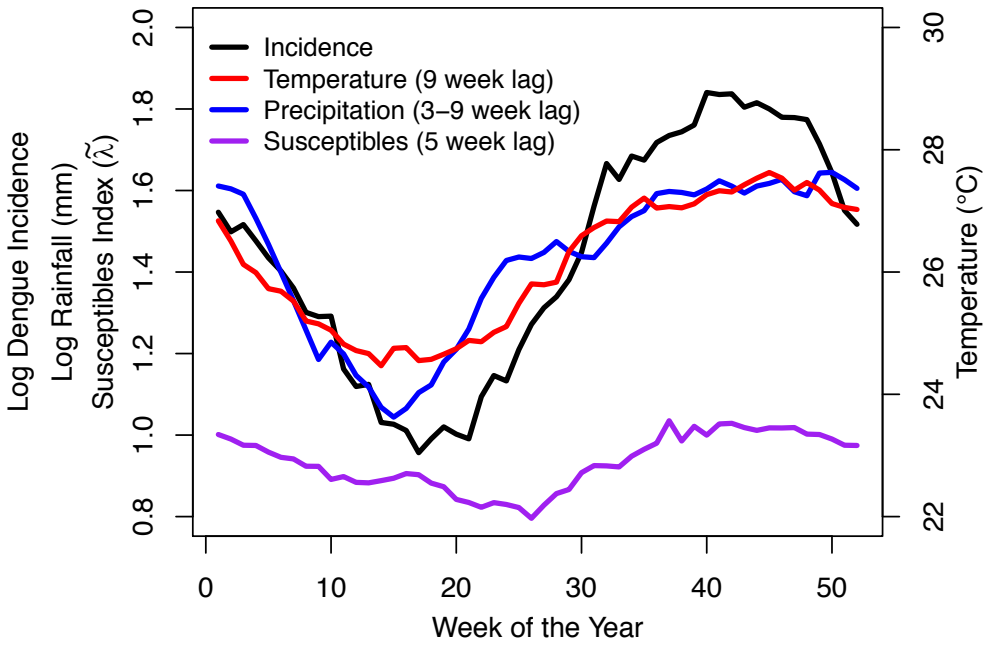


Figure 2. Seasonal trends and lags of dengue incidence and its drivers. The strong seasonal signal of dengue cases and other variables suggests potential causal lags between dengue incidence and temperature, rainfall, or proxy for the susceptible population size. The lines represent interannual averages for each week-of-year of dengue incidence (black), temperature lagged 9 weeks forward in time (red), average rainfall over the preceding 3–9 weeks and lagged 3 weeks forward in time (blue), and a proxy for susceptible population size lagged 5 weeks forward in time (purple).

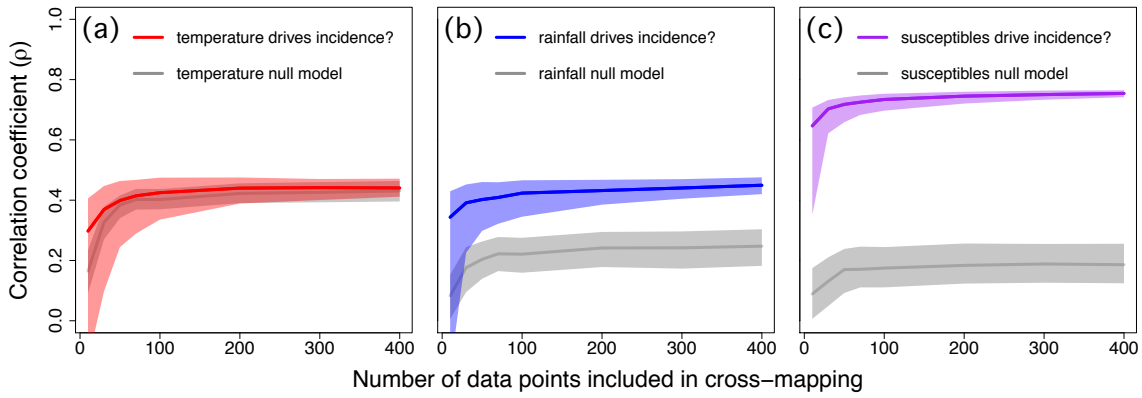


Figure 3. Susceptible population size and climate drive dengue incidence. Cross-mapping between dengue incidence and potential drivers—temperature with a 9-week lag (a; red), rainfall with an averaged 3–9-week lag (b; blue), and proxy for susceptible population size ($\tilde{\lambda}$) with a 5 week lag (c; purple)—display significant (Kendall's test $\tau > 0$; $p < 0.01$) convergence in cross-mapping skill (i.e., ρ increases and reaches a flat asymptote) as the number of time series data points increases (sign of causality). Cross-mapping skill is the Pearson's correlation coefficient, ρ, between predicted driver values using the univariable state space reconstruction of the response variable, and the observed driver values. Rainfall and susceptibles index showed significant forcing above and beyond seasonal signal ($p < 0.05$), because cross-mapping of the true time series (blue and purple) are distinguishable from their respective null models (grey). The red, blue and purple shaded regions represent the 0.025 and 0.975 quantiles of bootstrapped time series segments. The grey shaded regions represent the 0.025 and 0.975 quantiles of the seasonal null distributions obtained from 500 runs of randomized time series with conserved seasonal trends (Deyle *et al.* 2016a). The grey line represents the median of the null distribution.

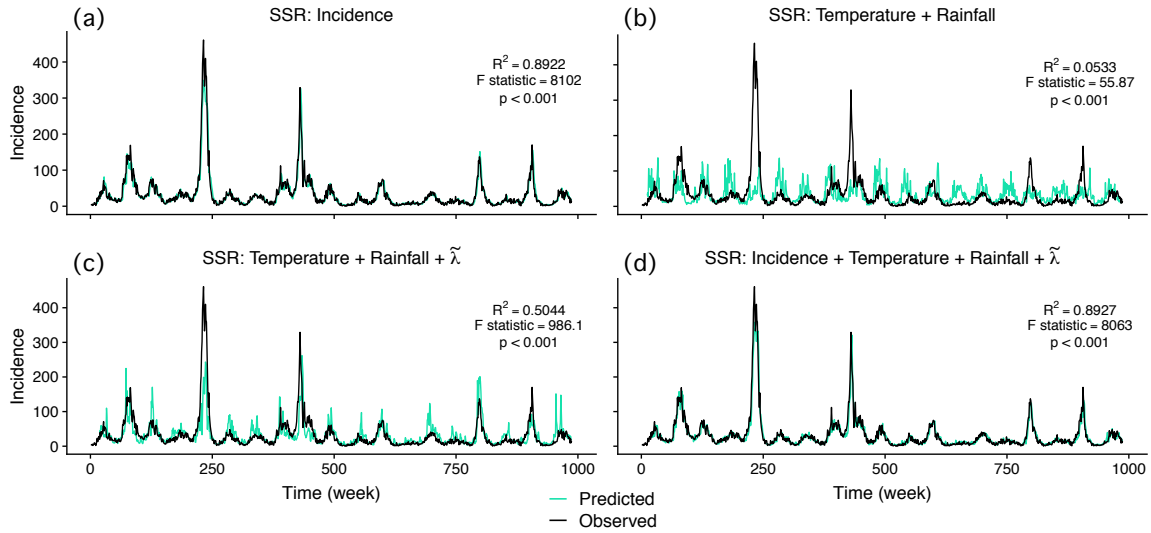


Figure 4. Predictive power of climate and susceptibles index ($\tilde{\lambda}$). Forecasting results showing observed (black) and predicted (green) values of dengue incidence (cases per week) using state space reconstruction (SSR). Univariable SSR with just incidence data (a) illustrates that dengue incidence is strongly predictable (R^2 = 0.8922, ρ = 0.9446). Multivariable SSR using only temperature and rainfall (b) predicts dengue incidence less well (R^2 = 0.0533, ρ = 0.2309), especially the magnitudes of the outbreaks, but the seasonal trend is captured. Multivariable SSR using temperature, rainfall, and the proxy for susceptible population size $\tilde{\lambda}$ (c) predicts incidence well (R^2 = 0.5044, ρ = 0.7102). Forecast predictions improve slightly (R^2 = 0.8927, ρ = 0.9448) when the drivers temperature, rainfall, and susceptibles index are added to incidence SSR (d). All R^2 values represent adjusted R^2. All SSR models (Figure 4a–d) had significant ($p < 0.001$) F statistics >> 1 based on ANOVA and the null hypothesis was rejected (H_0: there is no relationship between predicted and observed dengue incidence).

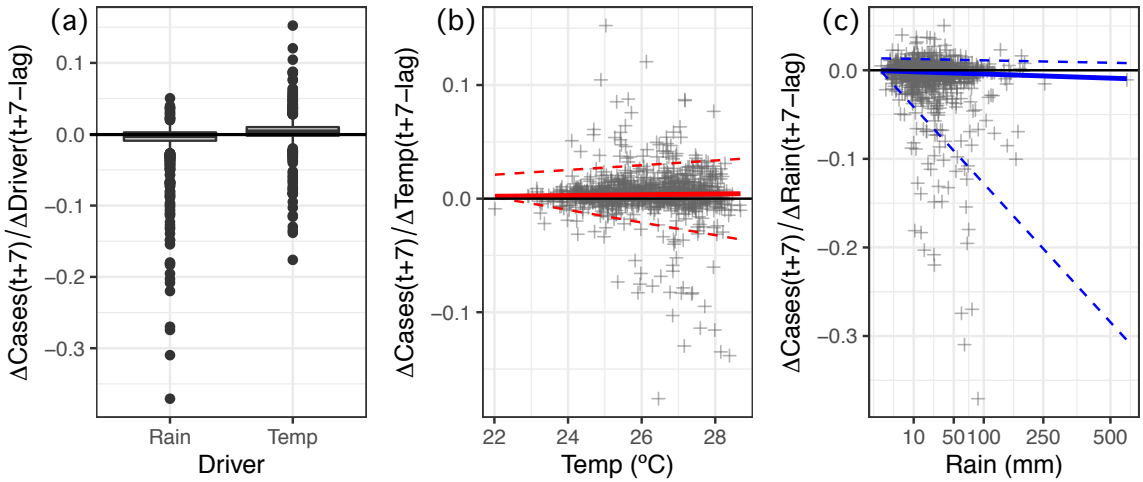
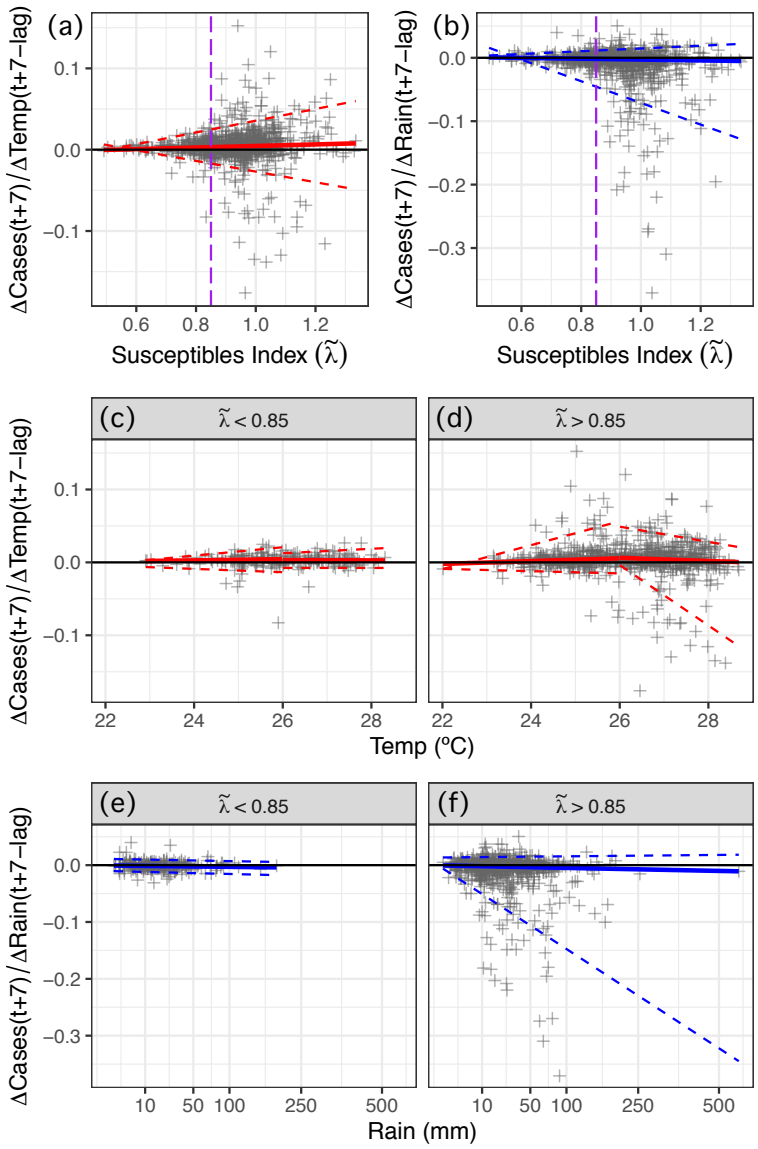


Figure 5. Temperature and rainfall show mixed functional responses of dengue incidence. Scenario exploration quantifies the variable effect of changes in drivers on dengue. Boxplots show that the median effects of rainfall (Rain) and temperature (Temp) are small (close to zero), but drivers occasionally have strong impacts (a). To investigate climate driver functional responses, we plotted the rate of change of dengue incidence as a function of temperature (b) and the rate of change of dengue incidence as a function of rainfall (c). Red and blue lines represent regression on the median for temperature and rainfall, respectively. The dashed red and blue lines represent regression on the 0.05 and 0.95 quantiles of temperature and rainfall, respectively. Temperature has an overall positive effect on dengue incidence (median regression line of the rate of change is positive), but can also have large negative and positive effects (a, b). Rainfall has an overall negative effect (median regression line of the rate of change is negative), but can also have small positive effects (a, c). Some of this effect variation can be explained by the susceptible population size (Figure 6).



710

711 **Figure 6. Temperature and rainfall effects on dengue vary depending on the**

712 **susceptible population size.** Effects of climate drivers (i.e., temperature and

713 rainfall) are investigated in relation to the proxy for susceptible population size, $\tilde{\lambda}$.

714 Plotting the effect of changes in temperature (a) and rainfall (b) against the

715 susceptible population size shows that driver effects are split around the threshold

716 of $\tilde{\lambda} = 0.85$ (purple dashed line). The red and blue lines represent the median

717 regression of temperature and rainfall effects, respectively (a–f). The dashed red

and blue lines represent the 0.5 and 0.95 quantile regressions of temperature and rainfall effects, respectively (a–f). Neither driver has an effect when there is not a sufficiently large susceptible population size ($\tilde{\lambda} < 0.85$) for an outbreak (c, e). However, above a certain susceptible population size the climate effects are observed: temperature with mostly a (possibly sigmoidal, see Figure S9) positive effect (d) and rainfall with a negative effect (f). However, even when driver effects are split at the evident threshold of $\tilde{\lambda} = 0.85$ (c–f), there are still many occurrences when the susceptible population size is sufficient large ($\tilde{\lambda} > 0.85$) but temperature and rainfall have no effect. In certain cases, temperature has even a negative effect on dengue (d).