Research

Cite this article: Hooshyar M, Wagner CE, Baker RE, Metcalf CJE, Grenfell BT, Porporato A. 2020 Cyclic epidemics and extreme outbreaks induced by hydro-climatic variability and memory. *J. R. Soc. Interface* **17**: 20200521. <http://dx.doi.org/10.1098/rsif.2020.0521>

Received: 2 July 2020

Accepted: 22 September 2020

Subject Category:

Life Sciences—Earth Science interface

Subject Areas:

environmental science

Keywords:

hydrology, infectious disease, epidemiology, dengue

Author for correspondence:

Amilcare Porporato

e-mail: aporpora@princeton.edu

Cyclic epidemics and extreme outbreaks induced by hydro-climatic variability and memory

Milad Hooshyar¹, Caroline E. Wagner⁴, Rachel E. Baker², C. Jessica E. Metcalf², Bryan T. Grenfell² and Amilcare Porporato³

¹CEE, PEI, and PIIRS, ²EEB and PEI, and ³CEE and PEI, Princeton University, Princeton, NJ, USA

⁴Department of Bioengineering, McGill University, Montreal, Canada

id MH, 0000-0002-3667-0482; CEW, 0000-0001-5193-2797; REB, 0000-0002-2661-8103; CJEM, 0000-0003-3166-7521

A minimalist model of ecohydrologic dynamics is coupled to the well-known susceptible–infected–recovered epidemiological model to explore hydro-climatic controls on infection dynamics and extreme outbreaks. The resulting HYSIR model reveals the existence of a noise-induced bifurcation producing oscillations in infection dynamics. Linearization of the governing equations allows for an analytic expression for the periodicity of infections in terms of both epidemiological (e.g. transmission and recovery rate) and hydrologic (i.e. soil moisture decay rate or memory) parameters. Numerical simulations of the full stochastic, nonlinear system show extreme outbreaks in response to particular combinations of hydro-climatic conditions, neither of which is extreme *per se*, rather than a single major climatic event. These combinations depend on the assumed functional relationship between the hydrologic variables and the transmission rate. Our results emphasize the importance of hydro-climatic history and system memory in evaluating the risk of severe outbreaks.

1. Introduction

Modelling the spread of infectious diseases is challenging not only because of the complexity involved in the mathematical representation of disease and human behaviour but also because the transmission of many infections is affected by the unpredictable and multiscale variability of hydroclimatic conditions. In several cases, the survival probabilities of disease agents (virus, bacterial and other parasites) are sensitive to hydro-climatic variables such as temperature, rainfall, air humidity [1], while in vector-borne diseases the life cycle of the vector is affected by climate and water availability [2–5]. A striking example of such a dependency is provided by dengue fever, whose vector mosquito (*Aedes aegypti*) thrives under specific temperature conditions and requires water availability to breed [6,7].

Hydro-climatic controls on disease spreading manifest themselves in specific temporal and spatial patterns of disease spread [8–11]. While these interactions between disease dynamics and hydroclimatic forcing are well known qualitatively, quantitative studies have mainly focused on analysing infection time series to establish a connection between fundamental epidemiological characteristics and hydro-climatic variables such as temperature, humidity, precipitation and soil moisture [12,13]. Such empirical relationships can be extrapolated in time with proper climatic predictions to explore the burden of disease in the future [14].

The study of infectious disease is often rooted in minimalist models such as the susceptible–infected–recovered (SIR) model or various alternative formulations thereof [15–17]. Under proper transition dynamics at the individual level, such models accurately represent the mean-field dynamics (e.g. the density of infected individuals) in the limit of large population size [18,19].

Uncertainty in the dynamics of such systems may arise from intrinsic demographic fluctuations or stochastic environmental forcings [19,20]. In addition, disease spread may be modulated by deterministic seasonal (periodic) forcings which are often related to annual climatic cycles (e.g. rainfall seasonality) or large-scale climatic oscillations (e.g. El Niño) [10,21]. The treatment of environmental variability in epidemic (or population dynamic) models has been mostly done by considering uncorrelated (i.e. white) noise added to the epidemic parameters, a stochastic switching between a small number of environmental states, or periodic (seasonal) variation in model parameters [20,22].

Here, we employ a minimalist ecohydrologic model to describe the connection between climatic forcing and water availability in the soil (soil moisture) [23,24]. Soil moisture quantifies the amount of available water to plants, animals and microorganisms in the soil and near the surface. It is directly related to the abundance of water ponds as potential breeding sites for mosquitoes and is linked to hydrologic extremes such as drought and floods as well as to humidity conditions of the atmospheric surface layer and near-surface temperature [25]. Therefore, it is a relevant hydrologic variable for diseases whose transmission is known to be a function of water availability (e.g. dengue, cholera and malaria).

Soil moisture dynamics are governed by fundamental hydro-climatic characteristics such as rainfall intensity and frequency, atmospheric thermodynamic and meteorologic conditions, vegetation, soil type etc. [23,24]. In particular, the storage capacity of the top soil layer acts to dampen hydrologic variability, thus maintaining a ‘memory’ of the climatic forcing (e.g. the sequence of rainfall events). From a mathematical point of view, this makes hydro-climatic forcing a type of coloured noise, privileging variability at lower frequencies compared to high-frequency components [26]. The strength of this memory, and in turn its impact on disease dynamics, is controlled by soil properties (e.g. porosity and texture), the depth of the rooting zone and active soil layer, vegetation cover and climatic variables (e.g. temperature) that define the maximum evapotranspiration.

Broadly motivated by diseases whose transmission is influenced by water availability through biological effects on the infectious parasite or transmitting vector (e.g. dengue, yellow fever and cholera), here we study the effect of soil moisture memory, as modulated by hydrologic forcing, on the dynamic behaviour of epidemics. Although in reality a variety of more nuanced models accounting for the distinct immunological responses associated with different diseases may be required to adequately model disease-specific population-level transmission dynamics, we focus here on the effect of hydrologic forcing on the dynamics of acute immunizing infections by considering the archetypal SIR model coupled with an ecohydrological soil-moisture model through a functional dependence of the disease transmission rate on soil moisture. The probabilistic treatment of precipitation variability makes it possible to investigate the connection between rainfall regimes and soil moisture memory on the temporal distribution of infections. Our analysis reveals a link between the frequency of noise-induced oscillations in epidemics and the soil moisture decay rate. We further show that extreme outbreaks occur in response to a specific sequence of hydro-climatic forcing rather than a single extreme climatic event. The particular

characteristics of these outbreak-inducing sequences change depending on the form of the functional dependence of the transmission rate on soil moisture and rainfall distributions. The coupling with a hydrologic model, which takes into account the memory of climatic forcing, is vital for evaluating the probability of extreme outbreaks over a range of epidemiological parameters.

2. The HYSIR model

2.1. Stochastic soil moisture dynamics

The hydrologic part of the model is a stochastic differential equation representing the soil water balance, vertically averaged over the root zone of depth (z [mm]) in a horizontally homogeneous area [23,24]:

$$\frac{ds_m}{dt} = R(t) - ET(s_m) - LQ(s_m, t), \quad (2.1)$$

where s_m is the normalized soil moisture (s_m varies between zero and 1). R [d^{-1}], ET [d^{-1}] and LQ [d^{-1}] are the normalized rainfall rate, evapotranspiration and percolation plus runoff. All of these fluxes are normalized by the soil storage capacity w_0 [mm], which depends on soil porosity, texture and the rooting depth. Rainfall is assumed to be a marked Poisson process with frequency λ [d^{-1}] and exponentially distributed rainfall depths with mean α . Following [24], ET increases linearly with s_m and reaches a maximum value ρ at $s_m = 1$, i.e. $ET = \rho s_m$. Thus ρ is the maximum (or potential) normalized evapotranspiration ET_{max} . Assuming saturation-excess condition rainfall in excess of soil capacity leaves the system as LQ .

Figure 1*a,b* shows a realization of the rainfall and soil moisture time series from numerical simulation of equation (2.1). The rainfall signal consists of discrete pulses that make soil moisture jump randomly. Between the rainfall events, the soil moisture decays deterministically with the rate ρs_m .

The steady-state PDF of soil moisture can be obtained analytically in the form of a truncated gamma distribution [24], along with its mean and variance (see appendix A). The parameters in equation (2.1) embed the controls of soil, plant and climate characteristics on the soil water balance. An important quantity is the soil moisture integral time scale (area under the autocorrelation function), which is a measure of the soil moisture ‘memory’ given by ρ^{-1} [26]. Smaller ρ corresponds to slower decay of s_m , thus a longer soil moisture memory. The soil moisture memory and its impact on disease dynamics are directly related to soil, plant and climate variables that define the maximum evapotranspiration and soil water storage. For instance, soils with higher porosity tend to have a larger capacity to store water, thus correspond to smaller ρ and longer soil moisture memory. Maximum evapotranspiration, which is linearly proportional to ρ , increases by the energy flux available for the evapotranspiration and the capacity of the atmosphere to transport moisture [27].

We should note that although soil moisture s_m is physically bounded within $[0, 1]$, it may be beneficial to neglect the upper bound [28]. This choice allows for the quantification of water ponding and flooding that occurs at saturation with a single variable. In this way, s_m broadly represents water availability in the system in the form of soil moisture, ponding or flooding.

Figure 1. The hydrologically driven SIR model. (a,b) The time series of rainfall and soil moisture. The parameters used for these simulation are $\lambda^{-1} = 2.5$ d, $\alpha = 0.1$ and $\rho = 0.1$ d $^{-1}$. (c,d) The time series of susceptible and infected individuals where the transmission rate $B = \beta s_m + \beta/2$. The parameters used in this simulation are $\beta = 0.7$ d $^{-1}$, $\gamma = 1/14$ d $^{-1}$ and $\eta = 5.5 \times 10^{-5}$ d $^{-1}$ (or 0.02 y $^{-1}$).

2.2. Environmental controls on the susceptible–infected–recovered model

The dynamics of an immunizing infectious disease in a well-mixed population are commonly represented by the SIR model. Susceptible individuals may get infected and thus transition to the infected class and eventually move to the recovered class, which are immune to reinfection. Each class is subject to demographic dynamics such as death and birth. In the simplest form, the transition probability of susceptible to infected and from infected to recovered are $\beta n_I / N n_S$ and $\gamma n_I / N$, where n_S and n_I are the number of susceptible and infected individuals, N is population size, β is the transmission rate and γ is the recovery rate (inverse of the average infection period) [18]. If the infection is not fatal, one can assume a uniform death rate across the classes while births only contribute to the susceptible population. Under these assumptions and setting equal birth and death rates η , it can be shown that the ratio of average susceptible (S), infected (I) and recovered (R) individuals to the total population size (N) for large N is [18,29]

$$\frac{dS}{dt} = -B(s_m)IS - \eta S + \eta \quad \text{and} \quad \frac{dI}{dt} = B(s_m)IS - \eta I - \gamma I. \quad (2.2)$$

The assumption of equal birth and death rate implies a constant population where $S + I + R = 1$, allowing us to model the dynamics with only two equations. The dynamics of such a system may be subject to deterministic seasonal (periodic) forcings related to annual climatic cycles (e.g. rainfall seasonality) or large-scale climatic oscillations (e.g. El Niño) [10,21]. Under such seasonal forcings, the system

exhibits rich behaviours including period-doubling bifurcations and transitions to deterministic chaos [17,30]. In addition, the uncertainty in such systems may arise from intrinsic demographic fluctuations or stochastic environmental forcings [19,20]. Under the assumption of large population size, the SIR model with intrinsic noise is approximated as a system of Langevin equations which exhibits a characteristic amplification of infections [18,19].

The transmission rate B is assumed to be a function of soil moisture, i.e. $B = B(s_m)$, and s_m is forced by precipitation as given in equation (2.1). Thus, equations (2.1) and (2.2) form a system of stochastic ODEs to study infection dynamics under hydro-climatic forcing.

3. Noise-induced oscillations

An interesting feature of the HYSIR model is that, while with constant values of soil moisture the long-term solution is a stable point with values of I and S constant in time, when soil moisture is forced by a stochastic rainfall (figure 1a,b) the system exhibits cyclic behaviour as shown in the simulated time series of I and S in figure 1c,d. This is reminiscent of ‘noise-induced transitions’ in nonlinear systems with random forcing, where a wide range of behaviours, such as bi-modality of the steady-state solution and cyclic fluctuation, are brought about by noise [19,31,32].

To investigate the origin of this transition to cyclic behaviour analytically and link its frequency to the main ecohydrological parameters, we considered the behaviour of the system of equations (2.1) and (2.2) under conditions of small noise perturbations in rainfall. This allows us to

Figure 2. Power spectral density from numerical simulation of equations (2.1) and (2.2) and the linearized model in equation (3.2) for the time series of soil moisture (a) and infected individuals (b). For this simulation $B = \beta s_m + \beta/2$ and $\beta = 0.7 \text{ d}^{-1}$, $\eta = 5.5 \times 10^{-5} \text{ d}^{-1}$, $\gamma = 1/14 \text{ d}^{-1}$, $\lambda^{-1} = 1.05 \text{ d}$ and $\rho = 0.1 \text{ d}^{-1}$, and rainfall events with constant depth 7 mm d^{-1} .

obtain analytical expressions for the power spectral densities of $I(t)$ and $S(t)$. This analysis begins by noting that, in the absence of rainfall fluctuations $R(t) = \bar{R}$ and from equation (2.1), the average soil moisture is $\hat{s}_m = \bar{R}\rho^{-1}$, and in these conditions the HYSIR model has a non-trivial steady state solution (the trivial solution corresponds to the disease-free condition):

$$\hat{S} = \frac{\gamma + \eta}{B(\hat{s}_m)}, \quad \hat{I} = \frac{\eta}{\gamma + \eta} - \frac{\eta}{B(\hat{s}_m)}. \quad (3.1)$$

The system of equations (2.1) and (2.2) can be linearized around this solution to get

$$\left. \begin{aligned} \frac{dx}{dt} &= -\rho x + \zeta(t), \\ \frac{dy}{dt} &= a_{11}x + a_{12}y + a_{13}z, \\ \frac{dz}{dt} &= a_{21}x + a_{22}y + a_{23}z, \end{aligned} \right\} \quad (3.2)$$

and

where x , y and z are the deviation of s_m , S and I from the steady-state solution and ζ is the stochastic variation of R over its average. The dependence of these coefficients on the SIR and soil moisture parameters are reported in appendix B.

The behaviour of the model under rainfall perturbations can be studied in the frequency domain by taking the Fourier transform of equation (3.2). This results in the following analytical expression for the power spectral density (PSD) of soil moisture [26]:

$$\mathcal{P}_{s_m}(\omega) = \frac{\sigma_R^2}{\rho^2 + \omega^2}, \quad (3.3)$$

where σ_R^2 is the variance of rainfall. The PSD of soil moisture is defined by the parameter ρ and the total energy of the rainfall signal, σ_R^2 . At small wavenumbers, the soil moisture resembles a white noise process with a flat energy $\mathcal{P}_{s_m} \approx \sigma_R^2/\rho^2$. The soil moisture signal at high wavenumbers is that of red noise with $\mathcal{P}_{s_m} \propto \omega^{-2}$ [26], due to the low-pass filtering action of the soil water balance, which ‘adds memory’ compared to the white noise character of the rainfall forcing.

Similarly, the periodicity of epidemics can be studied in the frequency domain by computing the PSD of S and I , both of which can be expressed in similar forms as

modulations of the soil moisture PSD:

$$\mathcal{P}_{S,I}(\omega) = \mathcal{P}_{s_m}(\omega) \frac{A_{S,I} + B_{S,I}\omega^2}{C_{S,I} + D_{S,I}\omega^2 + \omega^4}. \quad (3.4)$$

The corresponding values of the coefficients are also reported for completeness in appendix B. Figure 2 shows the PSD for I and s_m from numerical simulation of equations (2.1) and (2.2) and the analytic PSDs in equation (3.4). In this simulation, a linear dependence between transition rate and soil moisture is assumed, i.e. $B = \beta s_m + \beta/2$.

Clearly, rainfall stochasticity gives rise to a dominant frequency in the time series of infected individuals. The distribution of the energy across frequencies is controlled by the parameters of the epidemic model as well as hydrologic memory. In the unrealistic limit of soil moisture becoming a white noise (e.g. no memory), the PSD of soil moisture is constant and appears as a coefficient (in place of $\mathcal{P}_{s_m}$) in equation (3.4). Thus with white noise the time series of infected individuals may still exhibit periodic behaviour. In general, however, the period of the oscillation depends on the degree of memory of the forcing, so the degree of the autocorrelation of the forcing is crucial. We also note that, in large but finite well-mixed populations, demographic noise is approximated as white additive noise and creates periodic behaviour [18,19].

Figure 3a,b shows the analytic PSDs of $I(t)$ and $S(t)$ from the linearized model for different ρ with constant disease-related and demographic parameters. The periodicity of $I(t)$ (inverse of the dominant frequency ω_m^{-1}) with different ρ and β is plotted in figure 3c. Smaller ρ results in a shorter periodicity in the infected time series, implying that faster soil moisture decay induces more frequent epidemics. From a hydrologic perspective, smaller ρ corresponds, in fact, to longer soil moisture memory, which is a characteristic of areas with lower maximum evapotranspiration or systems with high soil moisture capacity. The periodicity also decreases for higher transmission rates. Figure 3 shows the total power of the time series of infected individuals for a range of ρ and β . The total power is defined as $\int_0^\infty \mathcal{P}_I(\omega) d\omega$ which is equivalent to the variance of the infected time series σ_I^2 . Longer soil moisture memory (smaller ρ) increases the total power of fluctuations (i.e. total variance) in the time series of infected individuals, a trend that also occurs with decreasing β .

Figure 3. The power spectral density (PSD) of the time series of infected (a) and susceptible (b) individuals for a range of the hydrologic parameter ρ . For these results $\mathcal{B} = \beta s_m + \beta/2$ and $\beta = 0.7$ d $^{-1}$, $\eta = 5.5 \times 10^{-5}$ d $^{-1}$, $\gamma = 1/14$ d $^{-1}$ and $\bar{R} = 0.04$ d $^{-1}$. In (c,d), the periodicity of $I(t)$ (inverse of the dominant frequency ω_m^{-1}) and the total power of infected individuals time series σ_I^2 for a range of ρ and β are shown. The rest of the parameters are the same as those in (a,b). The PSDs are normalized by the variance of rainfall σ_R^2 .

Figure 4. The two generic functional forms $\mathcal{B}$ for the dependence of transmission rate on soil moisture. In Case I ($\mathcal{B} = \beta s_m + \beta/2$), transmission rate increases monotonically with stored moisture; whereas, Case II ($\mathcal{B} = 3\beta s_m(1 - s_m) + \beta/2$) represents an adverse effect on transmission rate for high moisture levels. These functions are chosen to have the same area under curve β for $0 \leq s_m \leq 1$.

4. Numerical simulations

The PSDs in equation (3.4) correspond to the conditions in which stochastic rainfall fluctuations are weak so that linearization is possible; this makes the results of the previous section only applicable to small perturbations. Hydrologic variability is seldom weak, however, and to explore the periodicity of epidemics more fully, including the magnitude of extreme outbreaks, in this section, we present the results of a comprehensive set of numerical simulations, in which the role of the memory of hydrologic forcing in relation to more realistic values of the main hydrologic variables is included.

To this end, we considered two generic functional forms to model the dependence of the transmission rate on soil moisture (Case I and II in figure 4). Case I corresponds to the condition where greater water abundance increases disease transmission. Case II, on the other hand, refers to a scenario in which increasing soil moisture initially results in increasing the transmission, but above a certain threshold transmission rate begins to fall. This situation may arise in vector-borne diseases such as dengue where high soil moisture can trigger flooding and destroy mosquitoes' breeding sites [33,34]. In these simulations, the birth/death rate was

Figure 5. (a–f) The period of infection ω_m^{-1} for different hydrologic parameters ρ , λ and $\bar{R}$ for both cases of $\mathcal{B}$ in figure 4. For these results $\beta = 0.7 \text{ d}^{-1}$ and $\gamma = 1/14 \text{ d}^{-1}$. The time series of infected individuals and their corresponding PSDs at points A and B are shown in (g–j).

set to $5.5 \times 10^{-5} \text{ d}^{-1}$ (or 0.02 y^{-1}) and each simulation was performed for $5 \times 10^6 \text{ d}$ with time step $dt = 1 \text{ d}$.

For the first set of simulations, we used the disease-related parameters $\beta = 0.7 \text{ d}^{-1}$ and $\gamma = 1/14 \text{ d}^{-1}$. For $s_m = 0.5$, these parameters correspond to a basic reproductive number ≈ 10 . As mentioned earlier, the dynamics of soil moisture are described by the parameters ρ , λ and $\bar{R}$. Here, we consider a range of hydrologic variables $\rho \in [0.05, 0.3] \text{ d}^{-1}$, $\lambda \in [0.15, 0.9] \text{ d}^{-1}$, and $\bar{R} = 0.03, 0.04$ and 0.05 d^{-1} . These choices of $\bar{R}$ roughly correspond to mean annual rainfalls of 1800, 2400 and 3000 mm, if typical values of soil porosity ($n = 0.55$) and top soil layer depth ($z = 300 \text{ mm}$) are assumed. The variation of infection periodicity for the range of hydrologic parameters is shown in figure 5. Interestingly, even with a constant $\bar{R}$ (i.e. constant mean annual rainfall) the periodicity varies significantly with changes in rainfall inter-arrival time (λ^{-1}) and soil moisture memory (ρ^{-1}) in both scenarios for $\mathcal{B}$.

Figure 6 shows the average ($\bar{I}$) and the standard deviation (σ_I) of infected individuals time series for a range of hydrologic parameters and $\bar{R} = 0.04 \text{ d}^{-1}$. To emphasize the effect of hydro-climatic forcing on extreme outbreak events, we also computed the change in I with a 10-year return period, i.e. an outbreak that is expected to occur every 10 years. In general, the changes in σ_I and I_{10y} are more pronounced than they are in $\bar{I}$.

5. Hydrologically driven extreme outbreaks

In the presented model, the incidence dynamics are driven by the hydrologic forcing through rainfall. Within this specific set-up, it is interesting to explore what combinations of hydrologic conditions are responsible for the emergence of extreme outbreaks. To understand the mechanism behind the amplification of hydrologically driven outbreaks, we looked at the hydrologic conditions prior to the top 5% of outbreaks defined as the local maxima of simulated incidence time series. Figure 7a,c shows examples of the average soil moisture before outbreaks. Here, δt denotes the time relative to the outbreak and $\bar{s}_m(t)$ is

$$\bar{s}_m(\delta t) = \frac{1}{J} \sum_j \frac{1}{\Delta t} \int_{t_j + \delta t}^{t_j + \delta t + \Delta t} s_m(u) du, \quad (5.1)$$

where j counts the top 5% of outbreaks and J is the number of these outbreaks. Δt is a time interval for averaging the soil moisture and is set to 30 d. Extreme outbreaks occur in response to a sequence of hydrologic conditions; specifically, they happen when a relatively long period of unfavourable hydrologic conditions for disease transmission is followed by a short period of favourable conditions. For Case I, this sequence translates to a long period of low soil moisture

Figure 6. The effect of hydrologic parameters ρ , λ and $\bar{R}$ on (a,d) the average and (b,e) the standard deviation of infected individuals time series, and (c,f) the extreme outbreaks with 10-year return period for both cases of $\mathcal{B}$ in figure 4. Here, $\beta = 0.7 \text{ d}^{-1}$, $\gamma = 1/14 \text{ d}^{-1}$ and $\bar{R} = 0.04 \text{ d}^{-1}$. The trends for other values of $\bar{R}$ are qualitatively similar and not shown here.

Figure 7. (a,c) Examples of hydrologic condition prior to the top 5% of outbreaks. The 30-day averaged s_m is denoted by $\bar{s}_m$ and δt is the time from the outbreaks. (b,d) Probability distribution of $\bar{s}_m$ which highlights the distinct hydrologic condition right before and 600 days prior to the outbreaks. The parameters used are $\rho = 0.1 \text{ d}^{-1}$, $\lambda = 2.5 \text{ d}^{-1}$, $\bar{R} = 0.04 \text{ d}^{-1}$ for Case I, and $\bar{R} = 0.05 \text{ d}^{-1}$ for Case II. The rest of the parameters are the same as those in figure 6.

(arid conditions) followed by a short period of high soil moisture (humid conditions).

Prior to extreme outbreaks in Case II, a period of low transmissivity (high soil moisture) is followed by soil moisture close to 0.5 (maximum transmission rate, as shown in

figure 4). It should be noted that a long period of low soil moisture followed by $s_m \approx 0.5$ in Case II can potentially create a severe outbreak. However, in the selected range of parameters the existence of such long low moisture periods is not likely.

Figure 8. The amplification caused by hydrologic memory for Case I (*a,b*) and Case II (*c,d*) quantified by $\Delta\sigma_I$ and $\Delta I_{10\text{ yr}}$, which are the change in the respective variables if the hydrologic memory is considered. The model parameters for these simulations are $\rho = 0.1\text{ d}^{-1}$ and $R = 0.04\text{ d}^{-1}$. The epidemiological parameters are similar to those in figure 6.

5.1. Amplification due to hydrologic memory

The hydrologic model considered here maintains some memory of past climatic forcing due to the deterministic decay of soil moisture occurring between rainfall events. Here, we explore how this memory affects the dynamics of the epidemic. To do so, we compare the results of numerical simulations of the coupled hydrologic and SIR model with the case where a white noise (i.e. with no memory) replaces the soil moisture signal. Given the parameters ρ , λ and $\bar{R}$, the mean and standard deviation of soil moisture are given by equation (A1), which can be used to generate white noise with the same statistics. Figure 8 shows the effect of hydrologic memory on σ_I and the outbreaks with a 10-year return period for Case I and Case II and different transmission (β) and recovery (γ) rates. Here, the amplification is measured by $\Delta\sigma_I$ and $\Delta I_{10\text{ yr}}$ is the change in the respective variables if hydrologic memory is considered compared to the case of forcing the system by white noise. Thus, positive values in figure 8, which are observed over a range of β and γ , represent amplifications due to hydrologic memory.

The soil moisture dynamics alter the autocorrelation functions of the infected and susceptible time series as shown in figure 9. The autocorrelation functions are given by $P(\tau) = \int_{\omega} e^{2\pi i \omega \tau} \mathcal{P}(\omega)$, where $\mathcal{P}$ is the PSD of the zero-mean and unit variance time series of infected or susceptible individuals. Soil moisture memory results in a sharper decline in the autocorrelation function of $I(t)$ (figure 9*a*)

when compared with the case forced by white noise. This is also reflected in the time series of infected individuals where longer periods between and amplitudes of epidemics are observed with the introduction of hydrologic memory. On the other hand, the autocorrelation function of $S(t)$ (figure 9*b*) exhibits a more gradual decay corresponding to susceptible build-up between outbreaks that have become less frequent (and more severe) with hydrologic memory.

6. An application to dengue incidence in Sri Lanka

In this section, we present an application of the proposed model to dengue incidence in Sri Lanka. Hooshyar *et al.* [28] analysed district-specific dengue incidence from 2011 to 2016 across Sri Lanka and reported a functional dependence between water availability w and transmission rate:

$$\ln \frac{\mathcal{B}}{\bar{\mathcal{B}}_d} = 0.38w - 0.18w^2. \quad (6.1)$$

Here, w is a hydrologic variable representing water availability; it follows similar dynamics as s_m in equation (2.1) but it is not bounded by 1. This choice allows us to capture the effects of ponding and floods at saturation via the single variable w . The district-specific constants $\bar{\mathcal{B}}_d$ were also found from regression analysis in [28] and $\bar{\mathcal{B}}_d = 2.1\text{ d}^{-1}$ for the district of Colombo (figure 10*g*). Figure 10*a* shows the normalized transmission rate as a function of water availability which belongs to a similar class as Case II in figure 4. The drop at

Figure 9. The autocorrelation function of the infected (a) and susceptible (b) time series when the system is forced by white noise and stochastic soil moisture with $\rho = 0.1 \text{ d}^{-1}$. (c,d) The corresponding time series of I and S . Here, we show the results for Case I in figure 4, although the behaviour in Case II is qualitatively similar. The model parameters for these simulations are $\rho = 0.1 \text{ d}^{-1}$ and $\bar{R} = 0.04 \text{ d}^{-1}$. The epidemiological parameters are the same as those in figure 6.

high values of w may be related to the negative effects of flash-floods on the mosquito life cycle [28,34].

Using these estimates, we use the HYSIR model for a preliminary exploration of the interaction of internal self-sustained oscillations, produced by random hydrologic forcing, with external seasonal rainfall forcing. Towards this goal, figure 10b shows the simulated incidence time series using daily rainfall data in the capital Colombo between 1980 and 2016 acquired from the climate hazards group infrared precipitation with stations (CHIRPS) [35] which were used to force the water availability assuming $\rho = 0.1 \text{ d}^{-1}$. The rainfall time series is intermittent due to the stochasticity of rain events and also seasonal because of the two dominant monsoon seasons in the southwest of Sri Lanka [36,37]. Figure 10c,e shows the seasonality of water availability and incidence, which are both driven by the seasonality of the input rainfall. We should note that the seasonality pattern in figure 10e closely resembles the biannual dengue incidence pattern observed in the southern regions of Sri Lanka [21,28].

We also considered a case without seasonality by generating a rainfall time series from a marked Poisson process fitted to real rainfall data. Removing rainfall seasonality obviously eliminates the seasonality in water availability and incidence, as shown in figure 10d,f, but still produces a noise-induced periodicity in the incidence time series, as depicted in figure 10b. For reference, the incidence time series with a constant transmission rate (equivalent to the mean transmission rate when forced by the real rainfall) is also shown in figure 10b which exhibits a spiral convergence towards the stable point.

While a more detailed analysis of external and internal periodicities is outside the scope of this paper and will be presented elsewhere, this preliminary analysis already highlights the importance of both seasonality and intermittency of hydro-climatic forcing for making an accurate prediction of disease dynamics. We note that lower frequency oscillations of climatic events such as El Nino also play an important role in water-related infections such as cholera [38]. The interaction of rainfall stochasticity, seasonality and low-frequency climatic events within our minimalist stochastic dynamical system is an interesting topic for further study.

7. Conclusion

Hydro-climatic variables such as temperature, rainfall, humidity and soil moisture are known to control various aspects of the dynamics of disease spread [1–4]. Soil moisture gives a quantitative measure of water availability near the surface and is relevant to the transmission of various pathogens such as dengue, cholera and malaria. Soil water availability may also be a relevant controlling factor in the dynamics of air humidity [39] through its connection to ground level humidity. Soil moisture dynamics include the input of rainfall and losses due to evapotranspiration, leakage and runoff; thus, they incorporate the fundamental hydro-climatic characteristics such as rainfall magnitude and distribution, vegetation cover and soil type [23,24].

Similarly to the amplification with intrinsic demographic noise [18], fluctuations caused by hydrologic uncertainty trigger a phase transition from a stable point to a cyclic solution with a characteristic periodicity in epidemics. The periodicity

Figure 10. (a) The relationship between transmission rate and water availability in Sri Lanka taken from [28]. Water availability follows similar dynamics to s_m in equation (2.1) but it is not bounded by 1. (b) The simulated incidence time series using real rainfall data from 1980 to 2016 in the capital Colombo (denoted by 'seasonal and intermittent'), reconstructed rainfall after removing seasonality (denoted by 'intermittent') and with a constant transmission rate. Simulations were initiated as $I = 0.001$ and $S = 0.03$. (c,e) The seasonality of water availability and incidence when the system is driven by a real rainfall signal. (d,f) The same graphs for rainfall without seasonality. The red lines in the background are the time series for each year and the black lines are the average of the ensembles of annual trends which denote the seasonality of the respective variable. The map of Sri Lanka is shown in (g). We assumed $\gamma = 1/14 \text{ d}^{-1}$ following [21] and $\eta = 5.5 \times 10^{-5} \text{ d}^{-1}$.

is mainly controlled by soil moisture memory, rainfall characteristics and epidemiological parameters. We observed that the periodicity varies significantly with changes in rainfall inter-arrival time. For instance, shorter periods between epidemics (figure 5b) are associated with more frequent rainfall, typical of the climate of tropics and southern Asia.

Our results suggest that if a dependence between water availability and disease transmission exists, the soil moisture memory dramatically affects the dominant frequency of epidemics and extreme outbreaks. The memory of soil moisture, which is a function of soil storage capacity and maximum evapotranspiration, is an essential component of the hydrologic system and allows for dampening of temporal variation in rainfall and maintaining ecosystem health. Lower $ET_{\max}$ corresponds to a slow decay of soil water content, and thus longer memory in the hydrologic system, a condition that is typical of temperate zones in mid-latitudes. The tropics are subject to relatively higher $ET_{\max}$, which corresponds to shorter soil moisture memory than that of the mid-latitudes.

The connection between extreme climatic conditions and disease outbreaks is another crucial aspect of human health, especially concerning the future risk of climate change [40]. Although such a relationship can be very complicated with several socio-economic and biological facets, our minimalist approach captures the coupling between extreme hydro-climatic conditions and disease outbreaks and highlights the importance of the risk of compound extremes.

Hydrologic controls on disease dynamics should also be studied in relation to the seasonality of climatic forcings. Future research will focus on the combined effects of seasonality and stochasticity of rainfall on the periodicity of infections and extreme outbreaks. This study aimed at understanding hydrologic impacts on diseases in which water availability affects the parasite or vector. However, other hydro-climatic variables such as temperature, rainfall, humidity may affect disease transmission through various mechanisms [1–4].

The spread of infections under hydro-climatic variation is a very complicated problem. The host and agent characteristics also modulate their responses to external forcings such as climate which further adds to the complexity of these dynamics. Although our analysis focused on acute immunizing infections, a wide range of immunological responses to different pathogens exist which may not be adequately captured by the SIR model. Such complexities include multi-strain interactions (e.g. dengue) and imperfect immunity (e.g. cholera), and extending our hydrologically forced framework to other epidemiological models is an important direction for future research.

Data accessibility. This article has no additional data.

Authors' contributions. M.H. performed the analysis and prepared the first draft which was edited first by A.P. and then all other authors. All authors contributed to the study design, discussion and interpreting results.

Competing interests. We declare we have no competing interest.

Funding. This work has been supported by the Princeton Environmental Institute and Princeton Institute for International and Regional Studies at Princeton University through the climate and disease initiative.

Acknowledgements. C.E.W. acknowledges support from the Life Sciences Research Foundation as an Open Philanthropy Project fellow. The numerical simulations in this article were performed on computational resources provided by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of

Information Technology's High-Performance Computing Center and Visualization Laboratory at Princeton University.

Appendix A

The mean and variance of equilibrium soil moisture distribution are

$$\bar{s}_m = \frac{\Gamma\left(\frac{\lambda + \rho}{\rho}\right) - \Gamma\left(\frac{\lambda + \rho}{\rho}, \gamma\right)}{\gamma\Gamma\left(\frac{\lambda}{\rho}\right) - \gamma\Gamma\left(\frac{\lambda}{\rho}, \gamma\right)} \quad \text{and} \quad \sigma_{s_m}^2 = \frac{\left[\Gamma\left(2 + \frac{\lambda}{\rho}\right) - \Gamma\left(2 + \frac{\lambda}{\rho}, \gamma\right)\right] \left[\Gamma\left(\frac{\lambda}{\rho}\right) - \Gamma\left(\frac{\lambda}{\rho}, \gamma\right)\right] - \left[\Gamma\left(1 + \frac{\lambda}{\rho}\right) - \Gamma\left(1 + \frac{\lambda}{\rho}, \gamma\right)\right]^2}{\gamma^2 \left[\Gamma\left(\frac{\lambda}{\rho}\right) - \Gamma\left(\frac{\lambda}{\rho}, \gamma\right)\right]}, \quad (\text{A } 1)$$

where $\Gamma()$ and $\Gamma(), \gamma$ are Euler and incomplete gamma functions.

Appendix B

The coefficient in equation (3.4) are computed from the Fourier transform of S and I in the linearized model and are given as

$$\left. \begin{aligned} A_I &= (a_{13}a_{21} - a_{11}a_{23})^2, \\ B_I &= a_{11}^2, \\ A_S &= (a_{12}a_{21} - a_{11}a_{22})^2, \\ B_S &= a_{21}^2, \\ C_I &= C_S = (a_{13}a_{22} - a_{12}a_{23})^2 \\ \text{and} \quad D_I &= D_S = (a_{12}^2 + 2a_{13}a_{22} + a_{23}^2). \end{aligned} \right\} \quad (\text{B } 1)$$

The coefficients a are

$$\left. \begin{aligned} a_{11} &= -\frac{dB}{ds_m}|_{\hat{s}_m} \hat{S}, \quad a_{12} = -B(\hat{s}_m) \hat{I} - \eta, \quad a_{13} = -B(\hat{s}_m) \hat{S} \\ \text{and} \quad a_{21} &= \frac{dB}{ds_m}|_{\hat{s}_m} \hat{I} \hat{S}, \quad a_{22} = B(\hat{s}_m) \hat{I}, \quad a_{23} = B(\hat{s}_m) \hat{S} - \eta - \gamma. \end{aligned} \right\} \quad (\text{B } 2)$$

References

- Lowen AC, Mubareka S, Steel J, Palese P. 2007 Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* **3**, e151. (doi:10.1371/journal.ppat.0030151)
- Rogers DJ, Randolph SE. 2006 Climate change and vector-borne diseases. *Adv. Parasitol.* **62**, 345–381.
- Yang HM, Macoris MdLdG, Galvani KC, Andrighetti MTM, Wanderley DMV. 2009 Assessing the effects of temperature on the population of *Aedes aegypti*, the vector of dengue. *Epidemiol. Infect.* **137**, 1188–1202. (doi:10.1017/S0950268809002040)
- Yamana TK, Eltahir EAB. 2013 Projected impacts of climate change on environmental suitability for malaria transmission in West Africa. *Environ. Health Perspect.* **121**, 1179–1186. (doi:10.1289/ehp.1206174)
- Yamana TK, Eltahir EAB. 2013 Incorporating the effects of humidity in a mechanistic model of *Anopheles gambiae* mosquito population dynamics in the Sahel region of Africa. *Parasites Vectors* **6**, 235. (doi:10.1186/1756-3305-6-235)
- Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW. 2011 Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **108**, 7460–7465. (doi:10.1073/pnas.1101377108)
- Moore CG, Cline BL, Ruiz-Tibén E, Lee D, Romney-Joseph H, Rivera-Correa E. 1978 *Aedes aegypti* in Puerto Rico: environmental determinants of larval abundance and relation to dengue virus transmission. *Am. J. Trop. Med. Hyg.* **27**, 1225–1231. (doi:10.4269/ajtmh.1978.27.1225)
- Rogers DJ. 2015 Dengue: recent past and future threats. *Phil. Trans. R. Soc. B* **370**, 20130562. (doi:10.1098/rstb.2013.0562)
- Bhatt S *et al.* 2013 The global distribution and burden of dengue. *Nature* **496**, 504–507. (doi:10.1038/nature12060)
- Pascual M, Rodó X, Ellner SP, Colwell R, Bouma MJ. 2000 Cholera dynamics and El Niño-southern oscillation. *Science* **289**, 1766–1769. (doi:10.1126/science.289.5485.1766)
- Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. 2016 Global environmental drivers of influenza. *Proc. Natl Acad. Sci. USA* **113**, 13081–13086. (doi:10.1073/pnas.1607747113)
- Ferrari MJ, Grais RF, Bharti N, Conlan AJK, Bjørnstad ON, Wolfson LJ, Guerin PJ, Djibo A, Grenfell BT. 2008 The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684. (doi:10.1038/nature06509)
- Atchison CJ, Tam CC, Hajat S, Van Pelt W, Cowden JM, Lopman BA. 2010 Temperature-dependent transmission of rotavirus in Great Britain and the Netherlands. *Proc. R. Soc. B* **277**, 933–942. (doi:10.1098/rspb.2009.1755)
- Baker RE, Mahmud AS, Wagner CE, Yang W, Pitzer VE, Viboud C, Vecchi GA, Metcalf CJE, Grenfell BT. 2019 Epidemic dynamics of respiratory syncytial virus in current and future climates. *Nat. Commun.* **10**, 1. (doi:10.1038/s41467-018-07882-8)
- Keeling MJ, Rohani P, Grenfell BT. 2001 Seasonally forced disease dynamics explored as switching between attractors. *Physica D* **148**, 317–335. (doi:10.1016/S0167-2789(00)00187-1)
- Grenfell BT, Bjørnstad ON, Finkenstädt BF. 2002 Dynamics of measles epidemics: scaling noise,

- determinism, and predictability with the TSIR model. *Ecol. Monogr.* **72**, 185–202. (doi:10.1890/0012-9615(2002)072[0185:DOMESN]2.0.CO;2)
17. Dalziel BD, Bjørnstad ON, van Panhuis WG, Burke DS, Metcalf CJE, Grenfell BT. 2016 Persistent chaos of measles epidemics in the prevaccination United States caused by a small change in seasonal transmission patterns. *PLoS Comput. Biol.* **12**, e1004655. (doi:10.1371/journal.pcbi.1004655)
 18. Alonso D, McKane AJ, Pascual M. 2007 Stochastic amplification in epidemics. *J. R. Soc. Interface* **4**, 575–582. (doi:10.1098/rsif.2006.0192)
 19. McKane AJ, Newman TJ. 2005 Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* **94**, 218102. (doi:10.1103/PhysRevLett.94.218102)
 20. Méndez V, Campos D, Horsthemke W. 2012 Stochastic fluctuations of the transmission rate in the susceptible-infected-susceptible epidemic model. *Phys. Rev. E* **86**, 011919. (doi:10.1103/PhysRevE.86.011919)
 21. Wagner CE, Hooshyar M, Baker RE, Yang W, Arinaminpathy N, Vecchi G, Metcalf CJE, Porporato A, Grenfell Bryan T. 2020 Climatological, virological and sociological drivers of current and projected dengue fever outbreak dynamics in Sri Lanka. *J. R. Soc. Interface* **17**, 20200075. (doi:10.1098/rsif.2020.0075)
 22. Hufton PG, Lin YT, Galla T, McKane AJ. 2016 Intrinsic noise in systems with switching environments. *Phys. Rev. E* **93**, 052119. (doi:10.1103/PhysRevE.93.052119)
 23. Rodríguez-Iturbe I, Porporato A, Ridolfi L, Isham V, Cox DR. 1999 Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation. *Proc. R. Soc. Lond. A* **455**, 3789–3805. (doi:10.1098/rspa.1999.0477)
 24. Porporato A, Daly E, Rodríguez-Iturbe I. 2004 Soil water balance and ecosystem response to climate change. *Am. Nat.* **164**, 625–632. (doi:10.1086/424970)
 25. Whan K, Zscheischler J, Orth R, Shongwe M, Rahimi M, Asare EO, Seneviratne SI. 2015 Impact of soil moisture on extreme maximum temperatures in Europe. *Weather Clim. Extremes* **9**, 57–67. (doi:10.1016/j.wace.2015.05.001)
 26. Katul GG, Porporato A, Daly E, Oishi AC, Kim H-S, Stoy PC, Juang J-Y, Siqueira MB. 2007 On the spectrum of soil moisture from hourly to interannual scales. *Water Resour. Res.* **43**, W05428. (doi:10.1029/2006WR005356)
 27. Monteith JL. 1965 Evaporation and environment. In *Symposia of the society for experimental biology*, vol. 19, pp. 205–234. Cambridge UK: Cambridge University Press.
 28. Hooshyar M, Wagner CE, Baker RE, Yang W, Vecchi G, Metcalf CJE, Grenfell BT, Porporato A. 2020 Dengue seasonality and non-monotonic response to moisture: a model-data analysis of Sri Lanka incidence from 2011 to 2016. *ESSOAr preprint*.
 29. McKane AJ, Newman TJ. 2004 Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E* **70**, 041902. (doi:10.1103/PhysRevE.70.041902)
 30. Keeling MJ, P Rohani 2011 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
 31. Horsthemke W. 1984 Noise induced transitions. In *Non-equilibrium dynamics in chemical systems*, pp. 150–160. New York, NY: Springer.
 32. Porporato A, D'Odorico P. 2004 Phase transitions driven by state-dependent Poisson noise. *Phys. Rev. Lett.* **92**, 110601. (doi:10.1103/PhysRevLett.92.110601)
 33. Fouque F, Carinci R, Gaborit P, Issaly J, Bicout DJ, Sabatier P. 2006 *Aedes aegypti* survival and dengue transmission patterns in French Guiana. *J. Vector Ecol.* **31**, 390–399. (doi:10.3376/1081-1710(2006)31[390:AASADT]2.0.CO;2)
 34. Lehman JA, Hinckley AF, Kniss KL, Nasci RS, Smith TL, Campbell GL, Hayes EB. 2007 Effect of Hurricane Katrina on arboviral disease transmission. *Emerg. Infect. Dis.* **13**, 1273. (doi:10.3201/eid1308.061570)
 35. Funk C *et al.* 2015 The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data* **2**, 1–21. (doi:10.1007/s40745-015-0036-x)
 36. Thambyahpillay G. 1954 The rainfall rhythm in Ceylon. *Univ. Ceylon Rev.* **12**, 224–273.
 37. Suppiah R. 1996 Spatial and temporal variations in the relationships between the southern oscillation phenomenon and the rainfall of Sri Lanka. *Int. J. Climatol.* **16**, 1391–1407. (doi:10.1002/(SICI)1097-0088(199612)16:12<1391::AID-JOC94>3.0.CO;2-X)
 38. Moore SM *et al.* 2017 El niño and the shifting geography of cholera in Africa. *Proc. Natl Acad. Sci. USA* **114**, 4436–4441. (doi:10.1073/pnas.1617218114)
 39. Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, Comrie A, Viboud C. 2013 Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS Pathog.* **9**, e1003194. (doi:10.1371/journal.ppat.1003194)
 40. Cann KF, Thomas DRh, Salmon RL, Wyn-Jones AP, Kay D. 2013 Extreme water-related weather events and waterborne disease. *Epidemiol. Infect.* **141**, 671–686. (doi:10.1017/S0950268812001653)