Supplemental Material (not copyedited or formatted) for: Colin H. Kyle, Jiawei Liu, Molly E. Gallagher. 2020. "Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen." The American Naturalist 195(3). DOI: https://doi.org/10.1086/707138.

Online Appendix for: Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen

Colin H. Kyle[1]
Jiawei Liu[1]
Molly E. Gallagher[1,2]
Vanja Dukic[3]
Greg Dwyer[1*]

1. Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637; 2. Current address: Department of Biology, Emory University, Atlanta, GA; 3. Department of Applied Mathematics, University of Colorado, Boulder, CO 80309; * Corresponding author; e-mail: gdwyer@uchicago.edu

Prepared using the suggested LaTeX template for *Am. Nat.*

Supplemental Material (not copyedited or formatted) for: Colin H. Kyle, Jiawei Liu, Molly E. Gallagher. 2020. "Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen." The American Naturalist 195(3). DOI: https://doi.org/10.1086/707138.

SIR Model with Demographic Stochasticity

Fig. 1 in the main text shows the predictions of a simple SIR model that allows for demographic stochasticity. In this model, transmission is density-dependent, so that the probability of a transmission event in a short time interval $(t, t + dt)$ is $\beta S I dt$, where S and I are the (integer) numbers of uninfected and infected hosts, while the probability of a removal event is $\mu I dt$. The deterministic prediction (pink line) uses the burnout equation (Keeling and Rohani, 2008), according to which the cumulative fraction infected i is the (implicit) solution of the equation $1 - i = \exp(-N_T(S(0)i + I(0)))$, where the threshold $N_T = \mu/\beta$.

The red lines, which enclose the central 90% of model predictions, were derived from an embedded jump chain, a recursion that provides an exact calculation of the probability of each possible outcome (Daley and Gani, 1999). The black circles show model simulations that we generated using a Gillespie algorithm, a type of event-based simulation algorithm (Keeling and Rohani, 2008), with 100 realizations at each initial host density.

A well known result for this model (Daley and Gani, 1999) states that the probability of no epidemic is equal to $(\mu/(\beta S(0) + \mu))^{I(0)}$. This probability falls rapidly towards zero as $I(0)$ increases, but it also falls as $S(0)$ increases. Both effects are clearly apparent in fig. 1.

Data Collection

To ensure that we were able to collect 100 larvae each week at each site, we first carried out qualitative surveys in the winter before each larval season, to roughly assess local densities. At each site that was identified as usable, we selected a square plot, 500m × 500m in area, as the location within which we would collect larvae after hatch. To quantify initial larval densities within these square areas, we carried out egg mass counts on five 1/40th hectare circles, which are known to provide reliable predictors of larval densities (Elkinton and Liebhold, 1990). One circle was placed in each corner of the collection square, and one was placed near the center of the collection square. The average of these 5 counts served as our estimate of the initial density

$S(0)$ for that site, in terms of egg masses per square meter. To convert to units of larvae per square meter, we assumed that 400 larvae hatch from each egg mass (Elkinton and Liebhold, 1990), but this assumption affects only the scaling of density, not the effects of density (Dwyer et al., 2000). To monitor weather conditions, we placed a weather station at the center of the square at each site, and we recorded rainfall, temperature and relative humidity. The weather station consisted of a RainWise tipping-bucket rain gauge connected to a RainWise RainLog data logger and a HOBO U23 Pro v2 External Temperature/Relative Humidity Data Logger - U23-002, which logged conditions every 5 minutes.

We collected larvae by capturing them in individual 2 oz plastic cups that contained artificial agar-based gypsy moth diet (Bell et al., 1981). Larvae were then reared in these cups in the laboratory at 21° C, a temperature that maximizes the chance that an *E. maimaiga* infection will successfully produce conidia (Hajek, 1999). Larvae were checked for death and signs of conidia twice a week for three weeks. Conidia are often visually apparent, but even when they are not, they can be easily seen at 400× under a light microscope, as can resting spores (Hajek, 1999). In any cases of uncertainty, we therefore examined smears from dead larvae under a microscope.

Larvae for experiments were hatched from eggs collected from a population in Roscommon County, Michigan, 10 kilometers from our collection sites in that area. To remove any baculovirus from the eggs, we disinfected them in a 4% by volume formalin solution, which eliminates baculovirus infections in hatchlings (Dwyer and Elkinton, 1995). To ensure that the lab-reared insects did not contact either the fungal and viral pathogens until they were deployed, we reared the insects on artificial diet in the laboratory at 25° C. Lab-reared insects were grown in a different room from field-exposed insects.

Experimental cages were made of aluminum screening, with dimensions 20 x 20 x 5 cm. Each cage was stocked with roughly 20 unexposed lab-reared larvae, and then was placed on the soil at the base of a tree. Because red oak (*Quercus rubra*) is a “most preferred” gypsy moth host tree species (Barbosa and Greenblatt, 1979), each tree was an overstory red oak located within each of the five egg-mass-count circles at each site. We placed two cages under each tree each

Supplemental Material (not copyedited or formatted) for: Colin H. Kyle, Jiawei Liu, Molly E. Gallagher. 2020. "Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen." The American Naturalist 195(3). DOI: https://doi.org/10.1086/707138.

week. One cage per tree was covered with a clear plastic box to reduce the exposure of larvae to air-borne conidia while the other was uncovered so that larvae were exposed to both conidia and resting spores, but the boxes had almost no effect. After 24 hours, larvae were removed to the laboratory, where they were reared individually on artificial diet at 21° C until pupation or death, in the same manner as larvae collected from feral populations.

Model Details

As we described in the main text, resting spores are only active for a few weeks during the larval period, but the mechanisms determining the beginning and end of this period are not well known (Hajek, 1999). In our field collections, however, we observed that the times of initiation and cessation of resting spore germination were positively correlated with latitude. Accordingly, because gypsy moth hatch time and development depend on the number of accumulated growing degree days above 10° C (Elkinton and Liebhold, 1990), which is in turn a function of latitude, we expressed the initiation and cessation times in units of degree days above 10° C. In the main text, we treat the initiation time as day $\tau = 0$, but for the purposes of estimation, we refer to the start time as $T_{R,\text{start}}$, and the stop time as $T_{R,\text{end}}$. Here R stands for resting spore, to distinguish these times from the time when conidia cease to be produced, which we symbolize as $T_{C,\text{end}}$.

Again as we described in the main text, infections in fifth and sixth instar larvae do not produce infectious conidia, instead producing resting spores for the next season (Hajek, 1997). To incorporate this phenomenon, we fit a parameter to estimate the number of degree-days above 10° C required for our host populations to reach the fifth instar, or $T_{\text{C,end}}$. For $\tau > T_{\text{C,end}}$, cadavers in the model cease producing new infectious conidia, although existing conidia may continue to cause infections.

Pseudo-Code for Numerical Algorithms

To explain our numerical algorithms in more detail, here we present the algorithms in the form of pseudo-code. Algorithm 1 shows pseudo-code for the numerical integration routine, while algorithm 2 shows pseudo-code for the line search MCMC fitting routine.

Algorithm 1 Numerical Integration Routine. Here we use the subscript τ to indicate the day during the epizootic. "RKF(4,5)" indicates a Runge-Kutta-Fehlberg routine of order 4,5.

```
function INTEGRATION(Observed Initial Host Density, Proposed Resting Spore Density)
    $S_0(0)$ = Observed Initial Host Density
    $R_{\tau \leq T}(0)$ = Proposed Resting Spore Density
    for $j$ = 1 to $m$ do $E_{0,j}(0) = 0$
    end for
    $C_0(0) = 0$
    for $\tau$ = 1 to Last Day of Epizootic do
        Draw random variates for stochastic terms for day $\tau$
        Calculate the transmission rates & the conidia decay rate from weather data for day $\tau$
        $S_\tau(0) = S_{\tau-1}(1)$
        for $j$ = 1 to $m$ do $E_{\tau,j}(0) = E_{\tau-1,j}(1)$
        end for
        $C_\tau(0) = C_{\tau-1}(1)$
        Integrate from $t = 0$ to $t = 1$, using RKF(4,5)
    end for
     return State Variables
end function
```

Likelihood Function

Because our data are binary, a binomial distribution could in principle be an appropriate likelihood function. The binomial distribution is based on the assumption that individual hosts are independent, but this assumption may be incorrect, for example if individuals that have higher or lower infection risk than average are clustered together within a population (McCullagh and Nelder, 1989). Any such lack of independence can cause the observed variance in infection risk to be substantially higher than the variance assumed under the binomial, a phenomenon known as over-dispersion. In the absence of direct information on the level of over-dispersion, a useful approach is to use a beta-binomial distribution (Cox and Snell, 1989), which is derived from the

Algorithm 2 Line-Search MCMC Algorithm. The algorithm consists of 3 major steps, which we depict here as 3 functions. In practice, line search and MCMC were implemented using purpose-built code in the `C` programming language, while the proposal generation routine uses a built-in PCA function implemented in the `R` programming language.

```
function LINE SEARCH(Data)
    for Iterations = 1 to 20 do
        for Parameters = 1 to Max Parameters  do
            for New Parameter Value = Lower Limit to Upper Limit do
                Calculate Likelihood Using New Parameter Value
                if Likelihood > Old Likelihood then
                    Keep New Parameter Value
                    Old Likelihood = Likelihood
                end if
            end for
        end for
    end for
     return Best Parameter Set
end function

function PROPOSAL GENERATION USING PCA(Line Search Results)
    Order line search results by likelihood
    Carry out PCA on top 10%
     return PCA Output
end function

function MCMC WITH AUTOMATED PROPOSALS(Data,PCA Output)
    for Iterations = 1 to Many do
        Propose New Parameters in PCA space
        Back-transform New Parameters onto original scale
        Calculate Likelihood using back-transformed New Parameters
        if Likelihood > Old Likelihood then
            Accept New Parameter
        else
            Draw U(0,1) Random Variate
            if Random Variate < (Likelihood/Old Likelihood) then
                Accept New Parameter
            else
                Discard New Parameter
            end if
        end if
    end for
end function
```

binomial by assuming that p, the probability of an infection in the binomial, follows a beta distribution, which describes quantities that vary between 0 and 1. In addition to the number of trials, which is equivalent to the sample size, n, the beta-binomial distribution has two parameters, as opposed to the single parameter of the binomial. Using a beta-binomial therefore allowed us to increase the variance as needed to explain the lack of fit of the model to the data, with the proviso that stochasticity can also help explain the lack of fit.

In practice, we defined the parameters of the beta-binomial to be a and b, and we re-parameterized according to:

$$a = pe^{\gamma}, \tag{S1}$$

$$b = (1-p)e^{\gamma}. \tag{S2}$$

Here p is again the probability of infection in the binomial, while γ is a parameter that affects the variance of p. For sample size n, under both the binomial and the beta-binomial, the expected number infected is np, but the variances of the two distributions differ. The variance of the binomial is $np(1-p)$, but the variance of the beta binomial is instead:

$$\text{Var} = np(1-p)\frac{n+e^{\gamma}}{1+e^{\gamma}}. \tag{S3}$$

This expression shows that, as $\gamma \to \infty$, the variance of the beta binomial approaches the variance of the binomial, but as $\gamma \to 0$, the variance instead approaches $np(1-p)(n+1)/2$, which is greater than the binomial variance for any $n > 1$, and equal to it for $n = 1$. By adjusting γ, we were thus able to adjust the variance of the beta-binomial as needed to explain the lack of fit between the model and the data. Because the models attempt to predict the value of p, γ is an inverse measure of the variance of the measurement process, and it thus quantifies the measurement error. Because we did not have independent information about the value of γ, we fit γ to the data, along with the other parameters. Because overdispersion may have been different for the wild-caught and experimental larvae, we included separate values of γ for wild-caught (γ_1) and experimental (γ_2) larvae.

Our estimates of the two γ values ranged from about 0.7 to about 2.2 (Tables S1-S3). These values are small enough to indicate that over-dispersion was likely occurring, but large enough to suggest that the level of over-dispersion was modest.

Including Observational and Experimental Data

Because we have both observational and experimental data, we calculated two likelihood scores for each parameter set, one for the observational data and one for the experimental data. Given a realization of the model, the two data sets are independent and can therefore be summed on a log scale. Because we expected that conditions in our experiments would be at least slightly different from conditions in nature, we included additional equations to describe the experimental data.

Because the experimental insects were only deployed in the field for 24 hours, none of them died of the disease, and so they did not contribute to the conidia population. Losses of conidia during the transmission process in the cages were also probably very small compared to the size of the conidia population. Meanwhile, our data consist only of the fraction infected, and so the length of time that the experimental insects spend in the exposed categories is of no interest. Also, because we expected that the infection rates would be at least moderately lower in the covered cages, we included separate equations for covered and uncovered cages:

$$\frac{dS_{c,\tau}}{dt} = -\theta_1 \nu_{C,\tau} S_{c,\tau} C_\tau - \theta_2 \nu_{R,\tau} S_{c,\tau} R_\tau(0) \tag{S4}$$

$$\frac{dS_{u,\tau}}{dt} = -\theta_3 \nu_{C,\tau} S_{u,\tau} C_\tau - \theta_4 \nu_{R,\tau} S_{u,\tau} R_\tau(0) \tag{S5}$$

Here $S_{c,\tau}$ is the density of insects in the covered cages on day τ, $S_{u,\tau}$ is the density of the insects in the uncovered cages on day τ, C_τ is the density of conidia on day τ, and $R_\tau(0)$ is the density of resting spores on day τ. These equations are thus connected to the full model equations through C_τ, which is a state variable in those equations. The crucial point is that θ_1 and θ_2 describe the difference in the conidial transmission rate $\nu_{C,\tau}$ and the resting spore transmission rate $\nu_{R,\tau}$ in the covered cages relative to the corresponding transmission rates in nature. The parameters θ_3 and θ_4 describe similar changes among the feral larvae. Because all larvae in the cages were fourth instars, and because the experimental period was quite short, for this model we did not include the degree-day growth function $D(\tau)$ in the transmission functions. Relative to the model in the main text, we are thus effectively assuming that fourth instars have $D(\tau) = 1$. Because the value of $D(\tau)$ for fourth instars in nature is of course much larger than 1, the values of θ take into account the effects on transmission of both differences in body size and differences in exposure conditions between experimental and feral larvae.

Algorithm Convergence

As we described in the main text, we tested for the convergence of our fitting algorithm using a Gelman-Rubin test (Gelman et al., 2014). To carry out this test, one must initiate multiple MCMC chains at different parameter values, such that the initial parameter values are proposed with inflated variances relative to the variances for the subsequent steps in the chain. The test statistic is then the ratio of a weighted sum of the average within- and between-chain variances to the average within-chain variance. Because the initial parameter values are proposed with inflated variances, the between-chain variance will always be larger than the average within-chain variances. As a result, the test statistic will always be larger than 1. Because of the way that the weighting is implemented, however, the test statistic approaches 1 as the between-chain variance approaches the within-chain variance.

In general, in the MCMC part of our line search/MCMC algorithm, we inflated the variances at the first step by multiplying them by 1.2. We then ran our chains until the Gelman-Rubin statistic fell below 1.1. In practice, most of our Gelman-Rubin statistics were below 1.01, and all were at or below 1.05. In our initial implementation of the MCMC step, however, we accidentally set the variance inflation factor to 0.5 instead of 1.2. Because of this mistake, we re-ran the MCMC step for the density-dependence only model, for the best weather only model, and for

Supplemental Material (not copyedited or formatted) for: Colin H. Kyle, Jiawei Liu, Molly E. Gallagher. 2020. "Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen." The American Naturalist 195(3). DOI: https://doi.org/10.1086/707138.

the better-fitting combined models, by which we mean all but the temperature-only and rainfall-only combined models. In all cases, we used an inflation factor of 1.2.

Because our line search-MCMC algorithm is reasonably robust, the median parameter values from this second set of MCMC chains were generally similar to the values from the first set, when the inflation factor was 0.5. The increased initial variance in the second set of chains led to a better exploration of parameter space, and so there were modest increases in parameter uncertainty for some parameters of some models. Because most of the models do not fit very well, however, their likelihood surfaces are probably rather flat, and so the potential for improvement in AIC scores was small. Because of this effect, for most models, AIC scores fell (improved) by only about 2 points when we used the second set of chains, for which the inflation factor was 1.2.

There are, however, exceptions to these trends. The AIC score for the density-dependence only model fell by 3.2, the score for the best model fell by 6.1, and the score for the second-best model fell by 12.1. The latter two changes likely occurred because those two models fit the data fairly well, and so their likelihood surfaces probably have sharp peaks. For the two best models, an increase in the initial variance therefore allowed the algorithm to find parameters that produced substantially better likelihoods.

For the models that we re-ran we thus report results from the second set of chains, for which we used a variance inflation factor of 1.2. For the worse-fitting weather-only and combined models, however, we report results based on the first set of chains, and so the AIC scores for those models are probably biased upwards slightly. Our reason for not re-running those models was that the AIC scores of the other poorly fitting models changed by less than 2 points, and such a small change would have had a negligible effect on AIC differences for the worse-fitting weather-only and combined models. Any bias in the AIC scores for those models would therefore almost certainly be too small to have changed our conclusions.

The Second-Best Model: Relative Humidity Only

As we explain in the main text, the second-best model has an AIC difference of 2.54, indicating that there is at least weak support for the second-best model. In fig. S1, however, we present the predictions of the second-best model, to show that, for that model, the fit to the data is noticeably worse than for the best model, with respect to both the epizootic data and the experimental data.

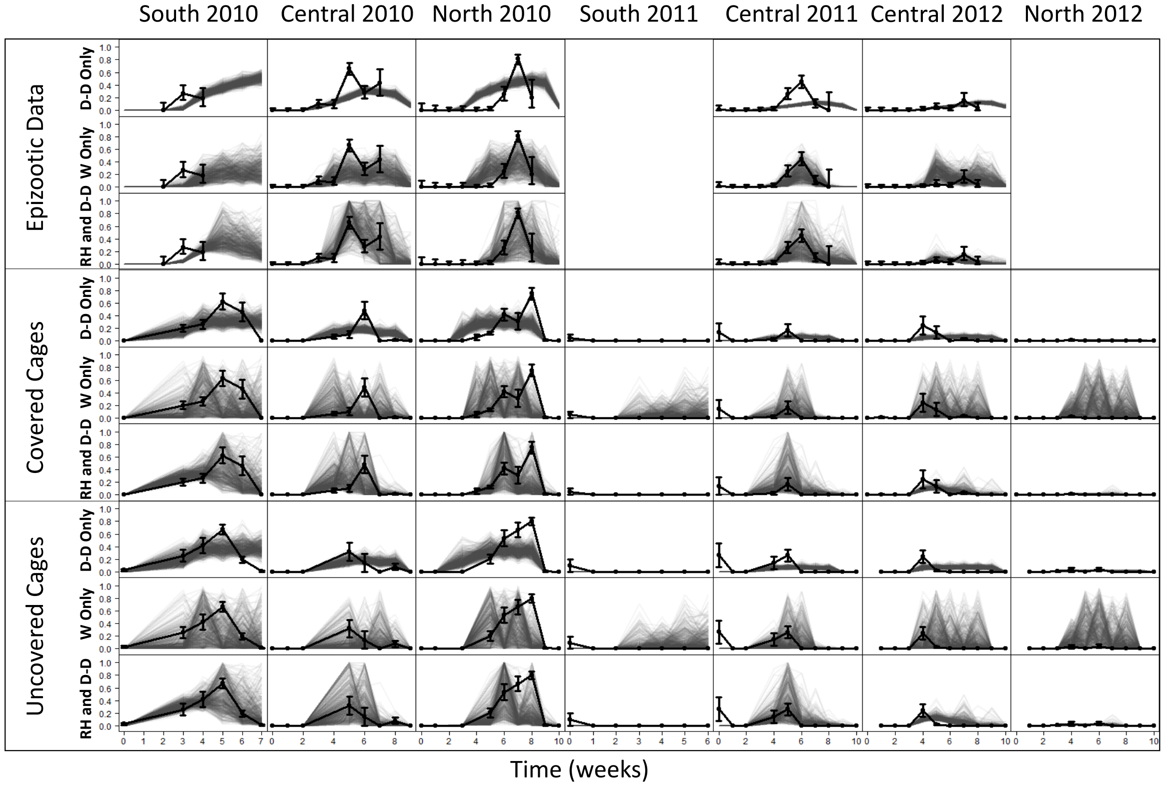

Figure S1: This figure is analogous to fig. 6 in the main text, which compares the fit to the data of the density-dependence only model, the best weather-only model, and the best model. In this case, however, we have replaced the predictions of the best model with the predictions of the second-best model, which includes effects of density-dependence and relative humidity, whereas the best model also includes the effects of rainfall and temperature. As in fig. 6 in the main text, here "D-D Only" refers to the density-dependence only model, "W Only" refers to the best weather only model, and "RH and D-D" refers to the model that includes the effects of density-dependence and relative humidity. Black solid lines are the data (+/- 1 SE) and the grey transparent lines are 25 realizations of the models using the best-fit parameter sets.

Moreover, the infection-rate curve for the second-best model is quite similar to the infection-rate curve for the best model (fig. S2). For the two models, the infection-rate curve increases at roughly the same rate with density, the threshold occurs at roughly the same density, and the spread in the density at which epizootics occurs is roughly the same (compare fig. S2 to fig. 7 in the main text). The only meaningful difference is that, for the second-best model, the infection rate at low density is substantially lower than it is for the best model.

A full understanding of why this latter difference occurs is beyond what we can accomplish here, but it seems likely that the explanation is as follows. Because the second-best model does not include rainfall, and because rainfall in our models affects only resting spore transmission, the second-best model cannot easily invoke high resting spore transmission to explain high infection rates. The best model in contrast does include effects of rainfall, and so, depending on rainfall, it can invoke high resting spore transmission to explain high infection rates. It thus seems likely that resting spore transmission in the second-best model plays less of a role than it does in the best model (the transmission functions for resting spores in the two models are so different that comparison of resting spore transmission parameters would be meaningless). Meanwhile, infections at low host density in any of our models must necessarily be due almost entirely to resting spores. This is because, when densities are low, there are too few hosts for there to be very many conidia. The lower infection rate at low host densities in the second-best model therefore appears to be due to that model's neglect of the effects of rainfall on resting spores.

These complications aside, the similarity of fig. S2 to fig. 7 in the main text confirms that our conclusions are robust to modest changes in model structure. That is, our main conclusion is that weather blurs but does not eliminate the disease-density threshold of *E. maimaiga*, and this conclusion holds for both the best model and the second-best model. Detailed comparison of the two models thus confirms our overall conclusions.

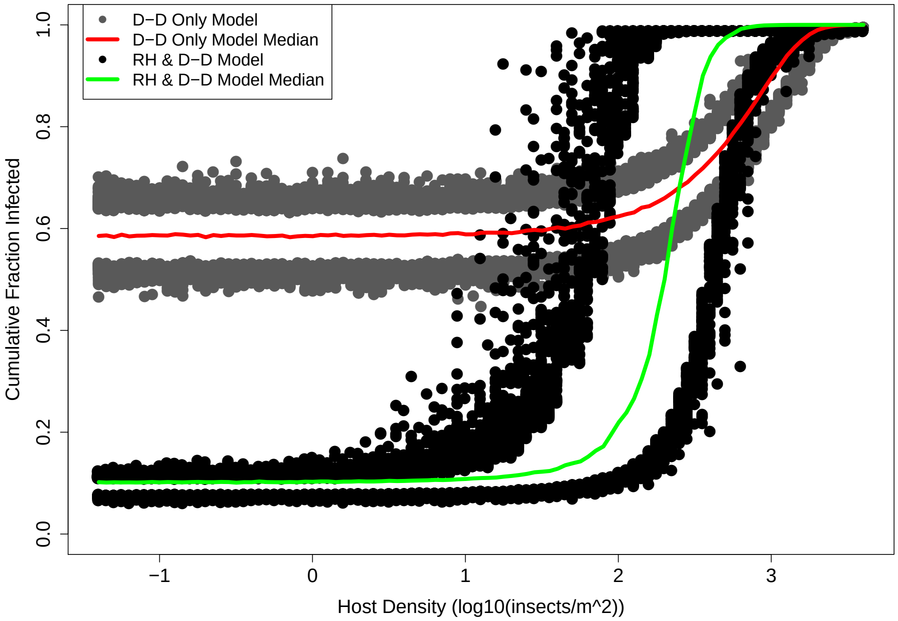

Figure S2: The predictions of the best density-dependence only model and the second best combined model, across a range of densities. As in fig. 7 in the main text, for each model, the points depict the upper 95% and lower 5% of 1000 realizations at each density, while the solid lines show the median at each density. For each simulation of the combined model, we generated weather using a Richardson weather generator (Richardson, 1981), which we fit using weather data from our study plots. The initial resting spore density for each model was taken from the respective model's best-fit value for the Central 2010 population. This density was lower for the density-dependence only model than for the best combined model (Online Appendix), but the density-dependence only model nevertheless has a high infection rate at low densities, due to its high resting spore transmission rate.

Linear models for weather variables

To build our weather generator, we compared the ability of different time-series models of varying complexity to explain the data logged by our weather stations at each field site (Burnham and Anderson, 2002). Initial tests showed that relative humidity and temperature are best predicted from rainfall. In our simulations, we therefore first used a log-normal distribution to describe rain $p(t)$ on day t, with a mean of 25.87 mm per day, and with variance 1189.2. We then generated values of maximum relative humidity $m(t)$, average temperature $\bar{h}(t)$ and maximum temperature $h(t)$ on day t using the following expressions:

$$\begin{aligned}
\bar{h}(t) &= 7.75 - 0.015 \times p(t) + 0.85 \times \bar{h}(t-1) \\
&\quad -0.18 \times h(t-2) - 0.016 \times p(t-4), \qquad \text{(S6)} \\
m(t) &= 17.0 + 0.210 \times p(t) + 0.064 \times p(t-1) \\
&\quad +0.052 \times p(t-4) + 0.54 \times m(t-1) \\
&\quad -0.73 \times \bar{h}(t) + 1.41 \times \bar{h}(t-1) \\
&\quad -0.43 \times \bar{h}(t-2), \qquad \text{(S7)} \\
h(t) &= 11.3 + 0.654 \times \bar{h}(t) - 0.011 \times R(t-4) \\
&\quad +0.25 \times h(t-1) - 0.054 \times h(t-2) \\
&\quad -0.094 \times m(t) \\
&\quad +0.023 \times m(t-1). \qquad \text{(S8)}
\end{aligned}$$

Parameter Estimates

Tables S1-S3 show the best-fit parameter values for the density-dependence only model, the best weather only model, and the best combined density-dependence plus weather model.

Part of the reason why we present the stochasticity terms in the main text as exponentiated, normally distributed, random variables is that the parameters of a normal distribution have more straightforward interpretations than the parameters of the corresponding log-normal distribution. To explain this, we first assume that, on a log scale, a stochasticity term is normally distributed with mean μ and standard deviation σ, as we assumed for the terms $\epsilon_{R,\tau}$, $\epsilon_{C,\tau}$, and $\epsilon_{F,\tau}$. On a non-log scale, the stochasticity term is then log-normally distributed, with mean $\exp(\mu + \sigma^2/2)$, and with variance $\big(\exp(\sigma^2) - 1\big)\exp(2\mu + \sigma^2)$. The variance expression in particular is sufficiently complicated that, in Tables S1-S3, we instead list the standard deviation parameters σ_R, σ_C, and σ_F. The latter parameters are thus the standard deviations of the corresponding normal distributions, which is to say, the standard deviations of the transmission rates on a log scale.

Table S1: Parameter Values for Density-Dependence Only Model. “95% HPD” indicates the 95% highest posterior density, a Bayesian estimate of parameter uncertainty that is conceptually similar to a confidence interval (Gelman et al., 2014). The parameters γ_1 and γ_2 are over-dispersion parameters, as we explain in the section entitled “Likelihood Function”.

Parameter		Median	95% HPD
$\nu_{R,t}$			
	ψ_1	4.29×10^{-2}	$[3.93 \times 10^{-2}, 4.71 \times 10^{-2}]$
	σ_R	0.494	[0.490, 0.498]
$\nu_{C,t}$			
	ψ_2	6.54×10^{-6}	$[5.95 \times 10^{-6}, 7.22 \times 10^{-6}]$
	σ_C	0.203	[0.158, 0.256]
$\mu_{C,t}$		0.408	[0.318, 0.517]
λ		0.110	[0.0979, 0.124]
$T_{R,start}$		120	[119, 122]
$T_{R,end}$		472	[470, 475]
$T_{C,end}$		336	[334, 339]
γ_1		1.30	[1.17, 1.43]
γ_2		0.713	[0.663, 0.764]
θ_1		634	[507, 800]
θ_2		2290	[2076, 2550]
θ_3		1128	[967, 1306]
θ_4		1838	[1679, 2007]
$R(0)$			
	2010 South	4.27×10^{-3}	$[3.78 \times 10^{-3}, 4.85 \times 10^{-3}]$
	2010 Central	1.52×10^{-3}	$[1.43 \times 10^{-3}, 1.61 \times 10^{-3}]$
	2010 North	3.76×10^{-3}	$[3.55 \times 10^{-3}, 3.97 \times 10^{-3}]$
	2011 South	1.08×10^{-12}	$[1.05 \times 10^{-12}, 1.12 \times 10^{-12}]$
	2011 Central	5.52×10^{-4}	$[5.11 \times 10^{-4}, 6.01 \times 10^{-4}]$
	2012 Central	6.33×10^{-4}	$[6.00 \times 10^{-4}, 6.68 \times 10^{-4}]$
	2012 North	1.14×10^{-4}	$[9.03 \times 10^{-5}, 1.47 \times 10^{-4}]$

Supplemental Material (not copyedited or formatted) for: Colin H. Kyle, Jiawei Liu, Molly E. Gallagher. 2020. "Stochasticity and Infectious Disease Dynamics: Density and Weather Effects on a Fungal Insect Pathogen." The American Naturalist 195(3). DOI: https://doi.org/10.1086/707138.

Table S2: Parameter values for the best weather only model (Rain + Temp), as in Table S1.

Parameter		Median	95% HPD
$\nu_{F,t}$			
	ψ_F	5.67×10^{-3}	$[4.88 \times 10^{-3}, 6.61 \times 10^{-3}]$
	ψ_4	2535	[621, 9218]
	ψ_5	0.138	[0.111, 0.172]
	ψ_9	0.247	[0.235, 0.260]
	σ_F	0.544	[0.532, 0.558]
λ		0.100	[0.0880, 0.112]
$T_{R,start}$		107	[104, 110]
$T_{R,end}$		463	[457, 470]
γ_1		1.23	[0.893, 1.71]
γ_2		0.855	[0.792, 0.928]
θ_2		2782	[2345, 3291]
θ_4		2364	[2086, 2659]

Literature Cited

Barbosa, P., and J. Greenblatt. 1979. Suitability, digestibility and assimilation of various host plants of the gypsy moth *Lymantria dispar*. Oecologia 43:111–119.

Bell, R. A., C. D. Owens, M. Shapiro, and J. R. Tardif. 1981. The gypsy moth: integrated pest management, chap. Mass rearing and virus production, pages 599–655. USDA Technical Bulletin, Washington, DC.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. Springer, New York, USA.

Cox, D. R., and E. J. Snell. 1989. Analysis of binary data, vol. 32. Chapman & Hall/CRC Press. New York, NY.

Daley, D. J., and J. Gani. 1999. Epidemic modelling: an introduction. Cambridge U. Press, Cambridge.

Dwyer, G., J. Dushoff, J. S. Elkinton, and S. A. Levin. 2000. Pathogen-driven outbreaks in forest defoliators revisited: building models from experimental data. The American Naturalist 156:105–120.

Dwyer, G., and J. S. Elkinton. 1995. Host dispersal and the spatial spread of insect pathogens. Ecology 76:1262–1275.

Elkinton, J. S., and A. M. Liebhold. 1990. Population dynamics of gypsy moth in North America. Annual Review of Entomology 35:571–96.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2014. Bayesian data analysis, third edition. Chapman & Hall/CRC Press. New York, NY.

Hajek, A. 1997. *Entomophaga maimaiga* reproductive output is determined by the spore type initiating an infection. Mycological Research 8:971–974.

Hajek, A. E. 1999. Pathology and epizootiology of *Entomophaga maimaiga* infections in forest Lepidoptera. Microbiology and Molecular Biology Reviews 63:814–835.

Keeling, M. J., and P. Rohani. 2008. Modeling infectious diseases in humans and animals. 1st ed. Princeton University Press, Princeton, New Jersey, USA.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall London.

Richardson, R. W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Research 17:182–190.

Table S3: Parameter values for combined weather plus density-dependence model (Rain + RH + Temp), as in Table S1.

Parameter		Median	95% HPD
$\nu_{R,t}$			
	ψ_3	1.31×10^{-2}	$[1.07 \times 10^{-2}, 1.56 \times 10^{-2}]$
	ψ_4	3.55	[1.39, 6.73]
	ψ_5	0.167	[0.121, 0.219]
	σ_R	0.372	[0.250, 0.519]
$\nu_{C,t}$			
	ψ_6	8.28×10^{-7}	$[6.42 \times 10^{-7}, 1.05 \times 10^{-6}]$
	ψ_7	7.05×10^{-2}	$[6.63 \times 10^{-2}, 7.46 \times 10^{-2}]$
	σ_C	0.914	[0.848, 0.984]
$\mu_{C,t}$			
	ψ_8	9.62×10^{-3}	$[5.81 \times 10^{-3}, 1.48 \times 10^{-2}]$
	ψ_9	0.234	[0.221, 0.248]
λ		0.120	[0.106, 0.134]
$T_{R,start}$		100	[97.0, 104]
$T_{R,end}$		267	[263, 271]
$T_{C,end}$		525	[515, 535]
γ_1		2.22	[2.04, 2.41]
γ_2		0.945	[0.893, 0.998]
θ_1		2424	[2186, 2688]
θ_2		2297	[2014, 2607]
θ_3		2054	[1878, 2235]
θ_4		1858	[1609, 2118]
$R(0)$			
	2010 South	1.83×10^{-2}	$[1.45 \times 10^{-2}, 2.23 \times 10^{-2}]$
	2010 Central	5.71×10^{-3}	$[4.72 \times 10^{-3}, 6.90 \times 10^{-3}]$
	2010 North	1.87×10^{-3}	$[1.48 \times 10^{-3}, 2.31 \times 10^{-3}]$
	2011 South	1.08×10^{-12}	$[1.02 \times 10^{-12}, 1.15 \times 10^{-12}]$
	2011 Central	2.33×10^{-3}	$[1.96 \times 10^{-3}, 2.75 \times 10^{-3}]$
	2012 Central	4.39×10^{-3}	$[3.98 \times 10^{-3}, 4.86 \times 10^{-3}]$
	2012 North	6.91×10^{-4}	$[4.85 \times 10^{-4}, 9.36 \times 10^{-4}]$