PNAS
www.pnas.org

Supplementary Information for

The duration of travel impacts the spatial dynamics of infectious diseases

John R. Giles, Elizabeth zu Erbach-Schoenberg, Andrew J. Tatem, Lauren Gardner, Ottar N. Bjørnstad, C. Jessica E. Metcalf and Amy Wesolowski

John R. Giles.
E-mail: giles@jhu.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.1922663117

Supporting Information Text

Hierarchical Bayesian model to estimate trip duration decay

To model the variation in duration of stay for commuter trips across different routes of travel, we estimated $N_{\text{decay}}(y_{ij})$, which is the expected number of commuters making a trip of duration y when travelling from origin district i to destination district j. The model uses an exponential decay function based upon the time spent y in district j that is linked to observations of call data records through a Normal error distribution that has inverse variance. The inverse variance is scaled by parameter ν_{ij}, which varies across all $i \to j$ routes.

$$\begin{aligned} N_{\text{decay}}(y_{ij}) &\sim \text{Norm}(\eta_{ij}, \varepsilon_{ij}) \\ \eta_{ij} &= N_{0_{ij}} e^{-\lambda_{ij} y_{ij}} \\ \varepsilon_{ij} &= (1/y_{ij})^{\nu_{ij}} \end{aligned} \quad [1]$$

The intercept term ($N_{0_{ij}}$) is the observed number of trips at $y = 0$ for each $i \to j$ route. We estimated decay rate parameters (λ) in the model hierarchically at both the population- and route-level to facilitate comparison of decay rates across different route types and compensate for routes that have lower sample sizes. Where, λ_{ij} in Equation 1 is the decay rate at the route-level, which is estimated by small modifications (Δ) to the population-level hyperparameter λ' such that $\lambda_{ij} = \lambda' \cdot \Delta_{ij}^{\lambda'}$. The population-level hyperparameter λ' was given the uninformative prior of Unif$(0, 25)$ and $\Delta_{ij}^{\lambda'}$ is a scaling factor with the prior Gamma$(2, 1)$. Figure S2 shows the hierarchical model graph.

The decay model likelihood uses a Normal link function with inverse variance. Based on preliminary analyses, we found that high rates of decay along with a long tailed distribution of observed trip duration required a higher weight to be placed on lower y-values. Therefore, we defined the variance of the model likelihood as $(1/y_{ij})^{\nu_{ij}}$, which takes the inverse of each y_{ij} value scaled by the route-level parameter ν_{ij}. The ν_{ij} parameter is estimated with the uninformative prior Gamma(1, 1) such that ν_{ij} is expected to be near 0 (all inverse weights ~ 1) and can be increased by the MCMC algorithm as needed for improved fit.

Gravity model incorporating both trip counts and trip duration. Initial data exploration suggested that the distribution of trip distances may be dependent upon trip duration. We developed a formulation of the gravity model that accounts for this interdependence between trip quantity and trip duration by incorporating the trip duration decay parameter λ_{ij} into the dispersal kernel so that the probability of movement to destination j also depends on the duration of stay at destination j. Note that subscripts ij represent movement from origin i to destination j ($i \to j$).

$$\begin{aligned} m_{ij} &\sim \text{Pois}\big(\pi_{ij} N_i\big) \\ \pi_{ij} &= c_{ij} / \sum_{\forall j} c_{ij} \\ c_{ij} &= \theta \left(\frac{N_i^{\omega_1} N_j^{\omega_2}}{f(d_{ij} \mid \lambda_{ij})} \right) \end{aligned} \quad [2]$$

Gravity model parameters are fit through normalized connectivity values (π_{ij}) that ensure integer values in the Poission likelihood function are scaled proportional to observed trip counts (m_{ij}). The exponential parameters ω_1 and ω_2 are weights that scale the contribution of origin and destination population sizes to the numerator, and θ is a proportionality constant. The denominator of the gravity model, $f(d_{ij} \mid \lambda_{ij})$, is comprised of a dispersal kernel function conditioned on trip duration.

We derived the conditional dispersal kernel function using Bayes theorem:

$$f(d_{ij} \mid \lambda_{ij}) = \frac{f(\lambda_{ij} \mid d_{ij}) f(d_{ij})}{f(\lambda_{ij})} \propto f(\lambda_{ij} \mid d_{ij}) f(d_{ij}). \quad [3]$$

Where, $f(\lambda_{ij} \mid d_{ij})$ is the expected rate of decay in trip duration given the distance between districts i and j, and $f(d_{ij})$ is a typical distance-based spatial dispersal kernel. This formulation of the conditional dispersal kernel $f(d_{ij} \mid \lambda_{ij})$ acts as a penalty on connectivity values that is proportional to the probability of a trip of distance d_{ij} that emanates from origin i given the decay rate parameter λ_{ij}. Values of the decay rate parameter were supplied by the mean of the posterior distribution of λ_{ij} estimated by the trip duration decay model.

To estimate the conditional dispersal kernel terms derived in Equation 3, we modeled $f(d_{ij})$ using the typical exponentiated term d_{ij}^{γ}, and $f(\lambda_{ij} \mid d_{ij})$ as the complement of the Empirical Cumulative Distribution Function (ECDF) of λ_{ij} with origin-specific model fitting parameters α_i:

$$f(d_{ij} \mid \lambda_{ij}) = d_{ij}^{\gamma} \big(1 - \text{ECDF}(\lambda_{ij})^{\alpha_i}\big). \quad [4]$$

Simulating disease dynamics

Disease dynamics were simulated using a stochastic Time series Susceptible-Infected-Recovered (TSIR) model (1–5). We apply the TSIR framework to a meta-population structure that allows us to include connectivity due to seasonal commuting and trip duration, which were estimated in the previous sections. We begin by defining local epidemic intensity as the expected number

of new infections $\mathbb{E}[I_{j,t+1}]$ at location j and time step $t+1$ by building upon the previous definition of spatial force of infection in (6).

$$\mathbb{E}[I_{j,t+1}] = \frac{\beta S_{jt}(I_{jt} + \iota_{jt} + \kappa_{jt})^{\alpha}}{N_{jt}} \quad [5]$$

The epidemic process in Equation 5 relies on the movement of infected individuals to track spatial diffusion of the pathogen, and assumes frequency dependent transmission with an absence of demographic stochasticity. The exponent α (typically < 1) is included to allow for nonlinearities in transmission and the stability of endemic equilibrium (3, 7). It can also be interpreted as the extent of population substructure that limits homogeneous mixing within that location (5). Since we model spatial dynamics with the mobility of infectious individuals, the susceptible population can be straightforwardly defined as:

$$S_{j,t+1} = S_{jt} - I_{j,t+1}. \quad [6]$$

The ι_{jt} term is a Poisson random variable with a mean equal to m_{jt}, which we define as the number infected individuals migrating to destination j from all other locations at time step t. The ι_{jt} term is typically used to model the effect of transient infections that arrive in location j at time step t, and remain for one full epidemic generation. However, data on the duration of trips made along each ij route allows us to adjust the temporal contribution of Infected individuals from other districts to each time step of the simulation. This is accomplished by defining m_{jt} as the sum of the number of infectious individuals at each origin i at time step t scaled by three terms; the probability that an individual leaves district i ($\hat{\tau}_i$), the estimated probability of travel from i to j ($\hat{\pi}_{ij}$), and the probability that an individual remains in destination j for a full epidemic generation when travelling from i to j ($\hat{\rho}_{ij}$). Therefore, the contribution to local dynamics from Infected individuals in all other districts is shown in Equation 7.

$$\begin{aligned} \iota_{jt} &= \text{Poisson}(m_{jt}) \\ m_{jt} &= \sum_{\forall i \neq j} \left(\hat{\rho}_{ij} \hat{\pi}_{ij} \hat{\tau}_i I_{it} \right) \end{aligned} \quad [7]$$

In addition to the infectious individuals that visit district j in time t (ι_{jt}), there are also infectious individuals that remain in district j from previous time steps, which we include as κ_{jt} in Equation 5. The κ_{jt} term is calculated as the number of infectious individuals that have migrated to district j in a previous time step $\iota_{j,t-\delta}$, multiplied by the mean estimated decay rate in trip duration $e^{-\delta \bar{\lambda}_j}$ for destination j, and the mean probability of remaining in destination j for a full epidemic generation after δ generations have passed $\bar{\rho}_j$. The summation over all previous time steps gives the estimated mean number of remnant infectious individuals r_{jt} shown in Equation 8.

$$\begin{aligned} \kappa_{jt} &= \text{Poisson}(r_{jt}) \\ r_{jt} &= \bar{\rho}_j \sum_{\delta=1}^{t} \left(\iota_{j,t-\delta} e^{-\delta \bar{\lambda}_j} \right) \end{aligned} \quad [8]$$

Both ι_{jt} and κ_{jt} terms are discrete random variables with Poisson error, which makes the force of infection of the local disease dynamics doubly stochastic and dependent on observed patterns in the human mobility data through both the immediate immigration of infectious individuals and the amount of infectious individuals remaining from previous immigration events.

Probability of leaving origin. We modeled the probability of an individual leaving district i ($\hat{\tau}_i$) as a continuous random variable with Beta distributed error (see Equation 9). We parameterized the Beta distribution with shape parameters a_i and b_i, which we derived from the mean μ_i and variance σ_i^2 of the observed proportion of individuals that left the origin district i at time t (x_{it}). Where the index t represents each unique day in the trip duration data. The x_{it} term was calculated by dividing the total number of individuals leaving origin i at time t $\left(N_{it}^{\text{leave}}\right)$ by the total number of observed trips emanating from origin i at time t $\left(N_{it}^{\text{leave}} + N_{it}^{\text{stay}}\right)$.

$$\hat{\tau}_i \sim \text{Beta}(a_i, b_i)$$

$$a_i = \mu_i^2 \left(\frac{1 - \mu_i}{\sigma_i^2} - \frac{1}{\mu_i} \right) \quad \text{and} \quad b_i = a_i \left(\frac{1}{\mu_i} - 1 \right) \quad [9a]$$

$$\mathbb{E}\left[x_i\right] = \mu_i = \frac{1}{T} \sum_{t=1}^{T} x_{it} \quad [9b]$$

$$\mathbb{E}\Big[\big(x_i - \mu_i\big)^2\Big] = \sigma_i^2 = \frac{1}{T}\sum_{t=1}^{T}(x_{it} - \mu_i)^2 \quad [9c]$$

$$x_{it} = \frac{N_{it}^{\text{leave}}}{N_{it}^{\text{leave}} + N_{it}^{\text{stay}}} \quad [9d]$$

Probability of remaining for full epidemic generation. We incorporated residence time into the spatial force of infection $\phi_{j,t+1}$ by adjusting the number of infectious individuals that contribute to a time step by the expected probability that visitors will remain at destination j for the full epidemic generation when travelling along route ij, which we denote as ρ_{ij}.

$$\rho_{ij} = \Pr(\text{remaining full generation in destination } j \mid \text{generation time } g)$$

The $\hat{\rho}_{ij}$ and $\bar{\rho}_j$ terms in Equations 7 and 8 are random variables drawn from a Beta distribution that is parameterized by the observed proportion of individuals p_{ij} that remain at destination j for the full epidemic generation when travelling along route ij.

$$\hat{\rho}_{ij} \sim \text{Beta}(a_{ij}, b_{ij})$$

Since the probability $\hat{\rho}_{ij}$ depends on the length of the infecting pathogen's generation time, we define it in terms of the generation time g, where the notation $g(n)$ indicates the length of time in days of n epidemic generations. Therefore, the shape parameters a_{ij} and b_{ij} are parameterized according to the observed mean μ_{ij} and variance σ_{ij}^2 of $p_{ij,g(n)}$, which is defined as the empirical proportion of individuals that remained in destination j for the full epidemic generation after n generation intervals in the trip duration data. Based on preliminary data analysis, we found that the majority of the variation in $p_{ij,g(n)}$ occurs among spatial locations (districts), therefore we reduced $p_{ij,g(n)}$ into its route-level mean μ_{ij} and variance σ_{ij}^2 (see Equation 10).

$$a_{ij} = \mu_{ij}^2\left(\frac{1-\mu_{ij}}{\sigma_{ij}^2} - \frac{1}{\mu_{ij}}\right) \quad \text{and} \quad b_{ij} = a_{ij}\left(\frac{1}{\mu_{ij}} - 1\right) \quad [10a]$$

$$\mathbb{E}\big[p_{ij}\big] = \mu_{ij} = \frac{1}{N_g}\sum_{n=1}^{N_g} p_{ij,g(n)} \quad [10b]$$

$$\mathbb{E}\Big[\big(p_{ij} - \mu_{ij}\big)^2\Big] = \sigma_{ij}^2 = \frac{1}{N_g}\sum_{n=1}^{N_g}(p_{ij,g(n)} - \mu_{ij})^2 \quad [10c]$$

The value of $p_{ij,g(n)}$ is calculated as:

$$p_{ij,g(n)} = \frac{\left[\mathbf{x} \cdot \left(\frac{\mathbf{g} - (g(n-1)-1)}{g(n) - (g(n-1)-1)}\right)\right]}{\sum \mathbf{x}}, \quad [11]$$

where the vector $\mathbf{g} = \{g(n-1), \cdots, g(n)\}$ contains all of the time steps (days) that fall within the n^{th} epidemic generation observed in the data and $\mathbf{x} = \{x_{ij,g(n-1)}, \cdots, x_{ij,g(n)}\}$ contains the counts of trips made for the corresponding generation. Therefore, Equation 11 provides the empirical proportion of individuals that have travelled from origin i to destination j that remain in destination j for all of the n^{th} epidemic generation.

Spatial infection hazard. Following (4), we calculated the time varying spatial hazard based on connectivity of each destination j at time t as

$$h(j,t) = \frac{\beta S_{jt}\big(1 - \exp(-x_{jt}\sum_{\forall i \neq j}\hat{\rho}_{ij}\hat{\pi}_{ij}\hat{\tau}_i y_{it})\big)}{1/(1+\beta S_{jt})}. \quad [12]$$

The term x_{jt} gives the proportion susceptible in destination j at time step t $\left(S_{jt}/N_{jt}\right)$, and y_{it} gives the proportion of infectious individuals in each origin i at time step t $\left(I_{it}/N_{it}\right)$. We then calculated the probability density function (PDF) for the waiting time of each district over all time steps with

$$w(j,t) = h(j,T)\prod_{t=1}^{T-1} 1 - h(j,t).$$

To calculate the probability of importation $p(j,t)$, we used a simple linear combination of all N_{sim} simulated realizations of $w(j,t)$ and then integrated the aggregate PDF by normalizing over all T time steps (8). We then calculated the peak of the aggregate PDF along with its 50% and 95% highest posterior density (HPD) intervals.

$$p(j,t) = \frac{\sum_{n=1}^{N_{\text{sim}}} w_n(j,t)}{\sum_{n=1}^{N_{\text{sim}}}\sum_{t=1}^{T} w_n(j,t)}$$

Gravity model without trip duration. To assess the influence of incorporating trip duration into the gravity model and spatial hazard (Equations 2 and 12), we compared our new formulations to those with all terms related to trip duration removed. First, we reduced the gravity model to a more basic form by replacing the conditional spatial dispersal kernel $f(d_{ij} \mid \lambda_{ij})$ with the distance-based dispersal kernel $f(d_{ij}) = d_{ij}^{\gamma}$ in Equation 2.

$$\begin{aligned} \pi_{ij}^{*} &= c_{ij} / \sum_{\forall j} c_{ij} \\ c_{ij} &= \theta \left(\frac{N_i^{\omega_1} N_j^{\omega_2}}{d_{ij}^{\gamma}} \right) \end{aligned} \tag{13}$$

In this formulation, π_{ij}^{*} indicates the probability of travel from district i to j under the basic gravity model. Note that Equation 13 was not refitted to the call data records, rather it was simulated using the estimated parameters from the full model with trip duration included (θ, ω_1, ω_2, and γ). Second, we removed all terms in the TSIR model that depend on the trip duration decay rate $\hat{\lambda}_{ij}$ or the probability of remaining for a full epidemic generation $\hat{\rho}_{ij}$, which results in the following TSIR model:

$$\mathbb{E}[I_{j,t+1}] = \phi_{j,t+1} = \frac{\beta S_{jt} (I_{jt} + \iota_{jt})^{\alpha}}{N_{jt}}. \tag{14}$$

Where,

$$\begin{aligned} \iota_{jt} &= \text{Pois}(m_{jt}) \\ m_{jt} &= \sum_{\forall i \neq j} \left(\pi_{ij}^{*} I_{it} \right) \end{aligned} \tag{15}$$

and the time varying spatial hazard based on connectivity of each destination j at time t is:

$$h^{*}(j,t) = \frac{\beta S_{jt} \left(1 - \exp(-x_{jt} \sum_{\forall i \neq j} \pi_{ij}^{*} y_{it})\right)}{1/(1 + \beta S_{jt})}. \tag{16}$$

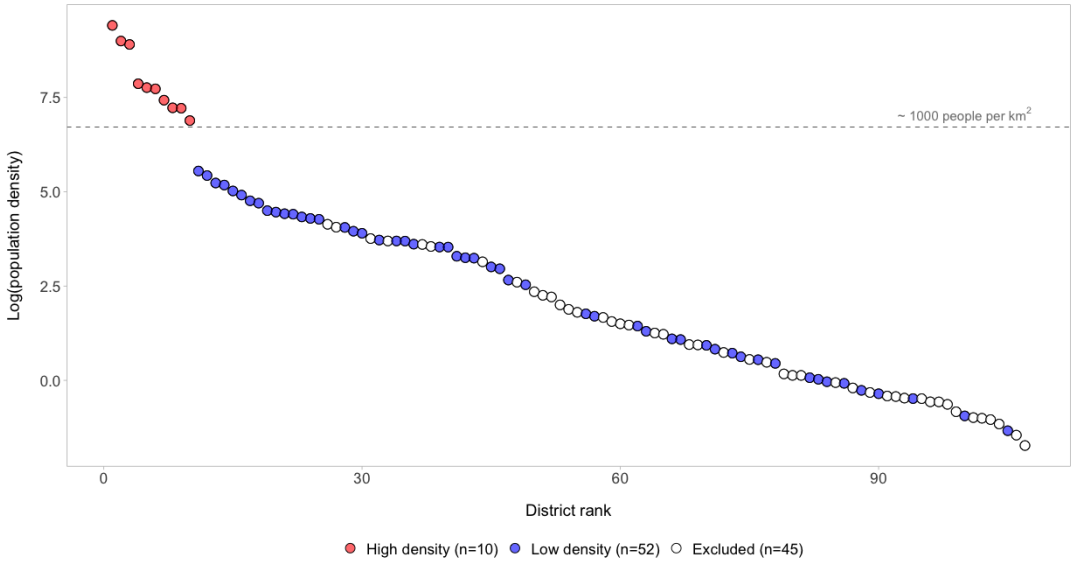

Fig. S1. The distribution of population density for districts included in analyses. Districts are ranked in order of district population density on the x-axis with log-transformed population density values on the y-axis. Districts with 'high' relative population density ($n = 10$) are shown in red, districts with 'low' relative population density ($n = 52$) in blue, and districts excluded due to low sample sizes ($n = 45$) are uncolored. The 'high' and 'low' density groups are defined using an arbitrary threshold of 980 people per km^2 that naturally delineates the 10 districts with noticeably higher population density than all other districts in Namibia (dashed line).

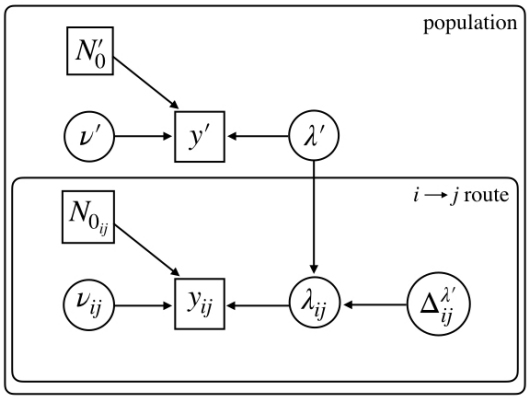

Fig. S2. Graph for the hierarchical trip duration decay model showing estimated parameters and scaling factors at the population- and route-level of the model. Parameters at the population-level are indicated by a prime symbol ($\prime$) and the scaling factor is indicated with a capital delta (Δ). Estimated parameters are shown in circular nodes and data or parameters derived from data are shown in square nodes.

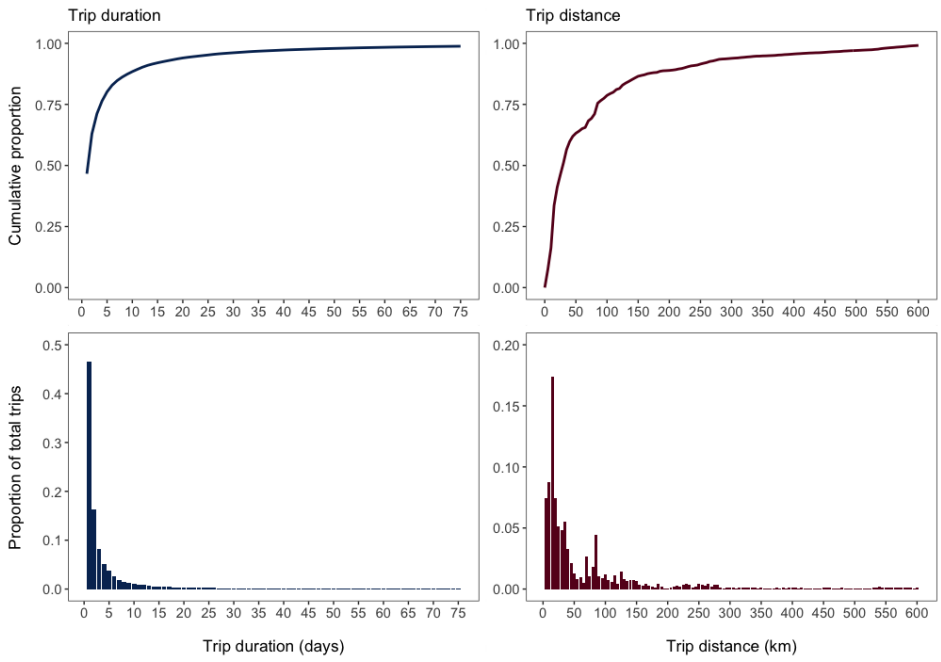

Fig. S3. The overall cumulative proportion and raw proportion of total trips for given values of trip duration and trip distance. See Table S1 for point measurements.

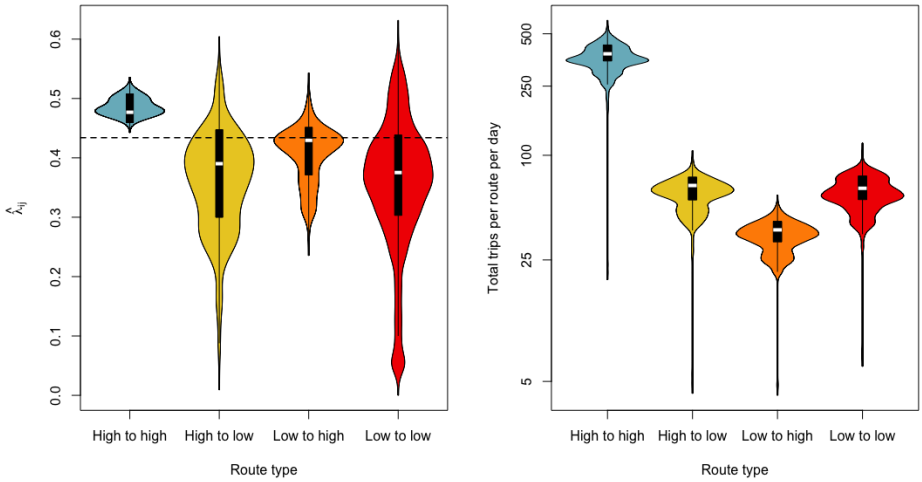

Fig. S4. Results from the Hierarchical Bayesian model that estimates the trip duration decay rate ($\hat{\lambda}_{ij}$). Values of $\hat{\lambda}_{ij}$ and proportion of total trips are plotted using a cutoff to define districts with 'high' or 'low' population density of 2500 people/km^2 (compared with 1000 people/km^2 in the main text). In panel A, the violin plots show the distribution of $\hat{\lambda}_{ij}$ for each of the four route-types compared to the population mean. In panel B, the violin plots show the distribution of the proportion of total trips for each of the four route-types over each day in the data set.

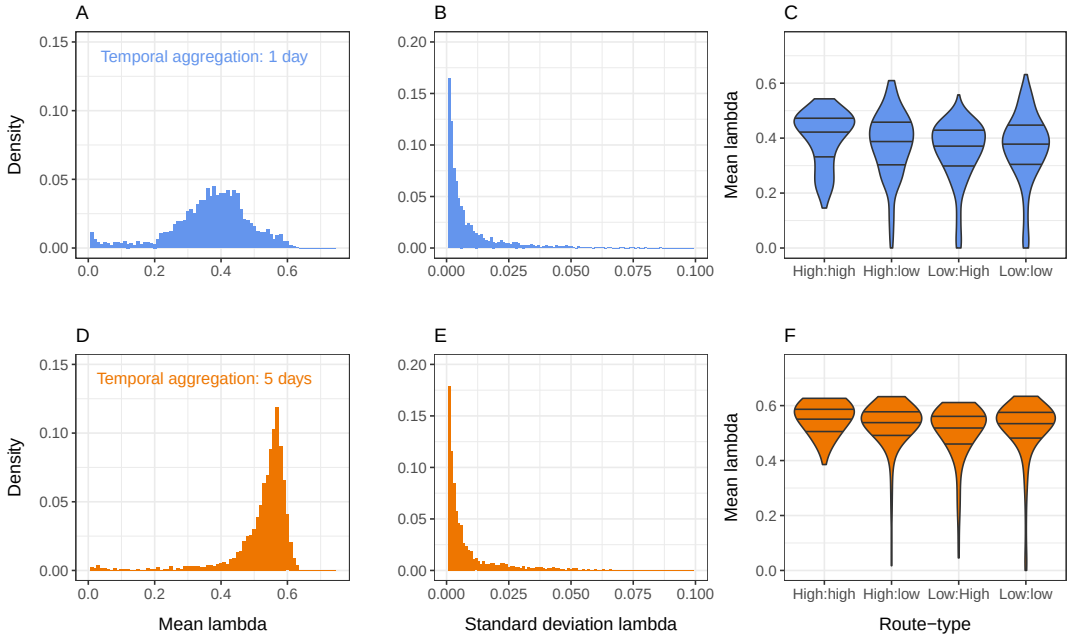

Fig. S5. Distribution of fitted trip duration decay rate parameters (λ_{ij}) where trip duration counts are aggregated to 1-day (A–C) and 5-day (D–F) temporal intervals. The first column shows the normalized density of all the posterior means of λ_{ij} (A and D) and the second column shows the standard deviation (B and E). The third column shows the distribution of mean λ_{ij} values for each route-type, indicating whether the origin and destination are categorized as high or low population density (C and F).

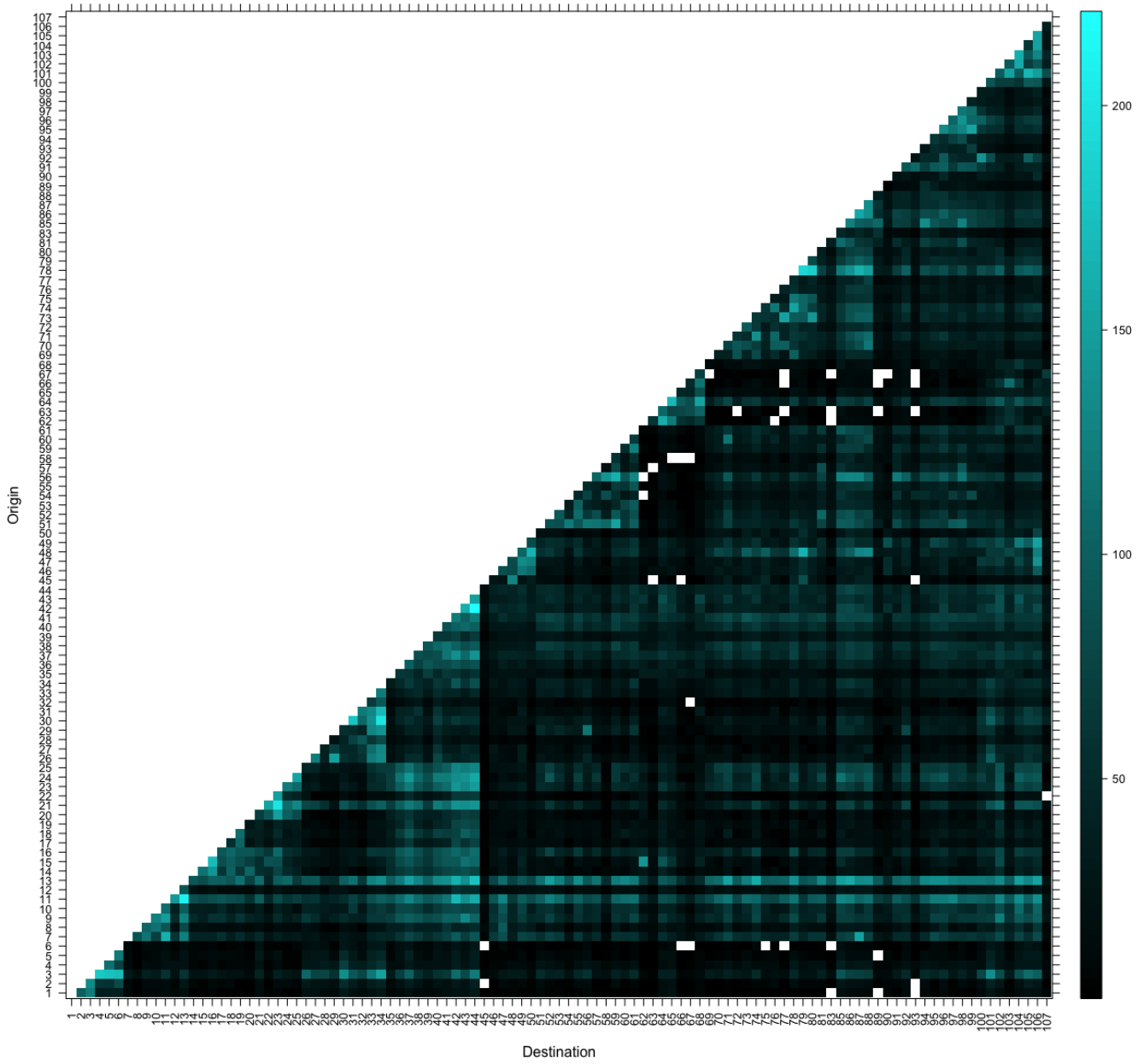

Fig. S6. Heatmap of sample sizes of observed trip duration counts for each ij route among all 107 districts.

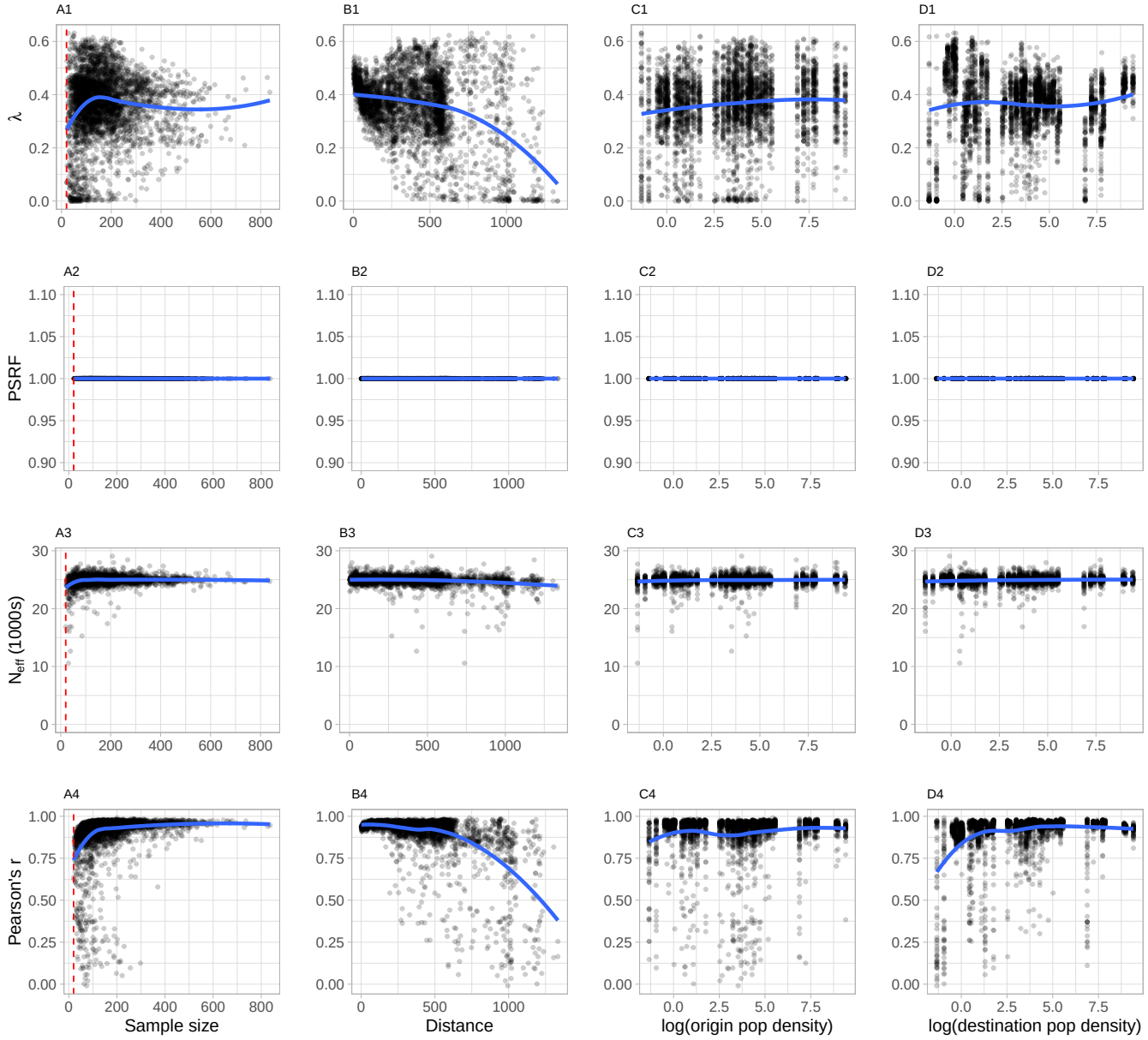

Fig. S7. Estimated trip duration decay parameter λ and model performance metrics plotted with: A) Sample size (number of unique observations of trip duration along each ij route), B) Distance between districts i and j, C) Population density (km^2) of origin district, D) Population density (km^2) of destination district. Decay parameter λ in row 1 show the mean of the posterior distribution of each route-level decay parameter (λ_{ij}). Model performance metrics are shown along rows 2–4: 2) Potential Scale Reduction Factor (PSRF) is a measure of model convergence with values near 1 indicating model convergence, 3) N_{eff} is the effective sample size of the posterior distributions for each λ_{ij}, and 4) Pearson's r gives the correlation between the fitted model response and observed trip duration counts. The dashed red line in the A plots indicates the minimum sample size of $n = 20$ unique observations of trip duration for all ij routes in the model.

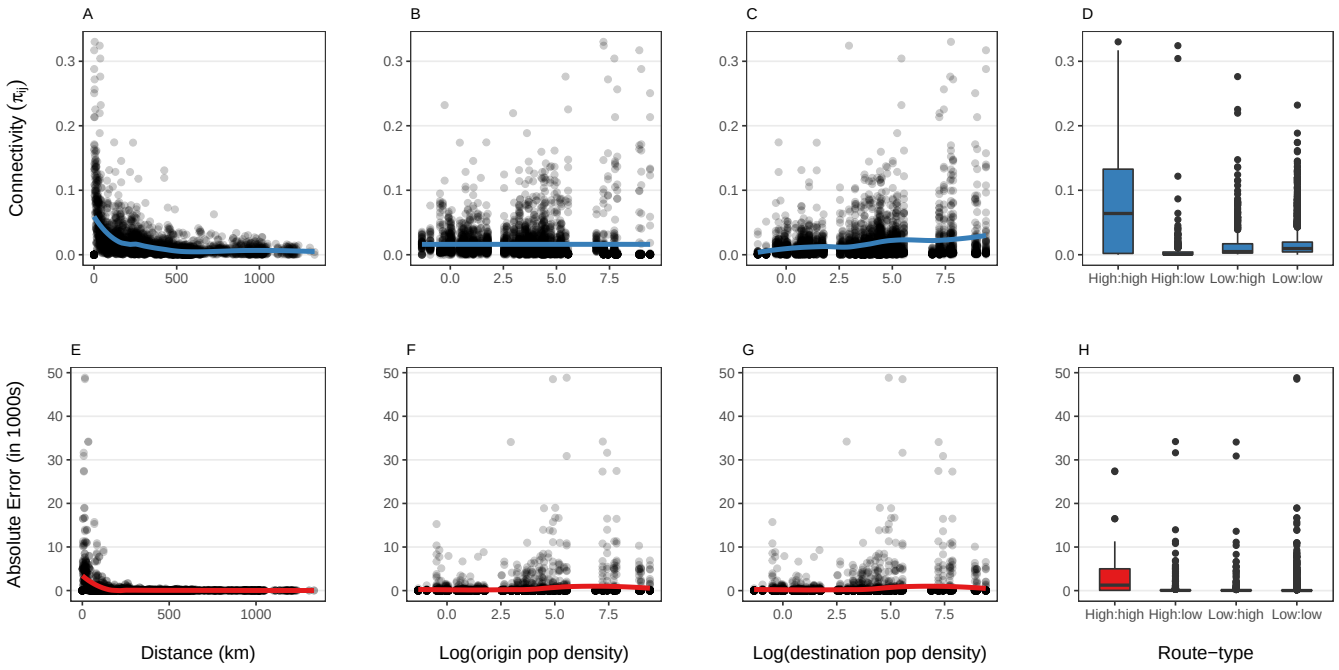

Fig. S8. Model performance of the trip duration decay model.

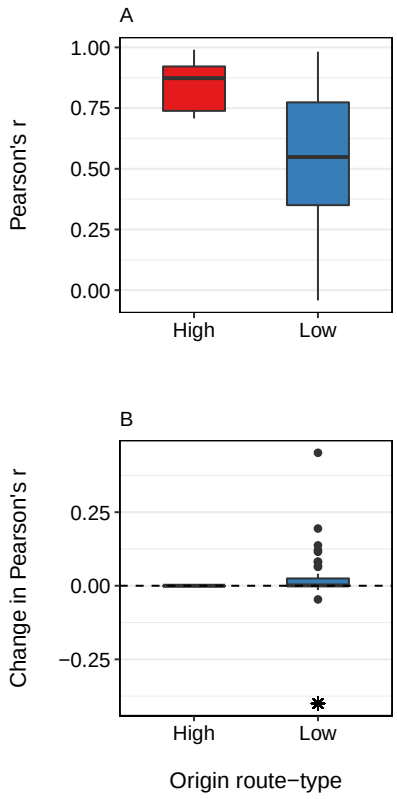

Fig. S9. Model performance of the trip duration decay model for routes that emanate from districts with either high or low population density. Panel A shows the distribution of Pearson's r of the basic gravity model for both rout-types and panel B shows the change in these values from for the duration gravity model. In panel B, the asterisk indicates district 24 (Luderitz), which was an outlier with a drastic decrease in model fit.

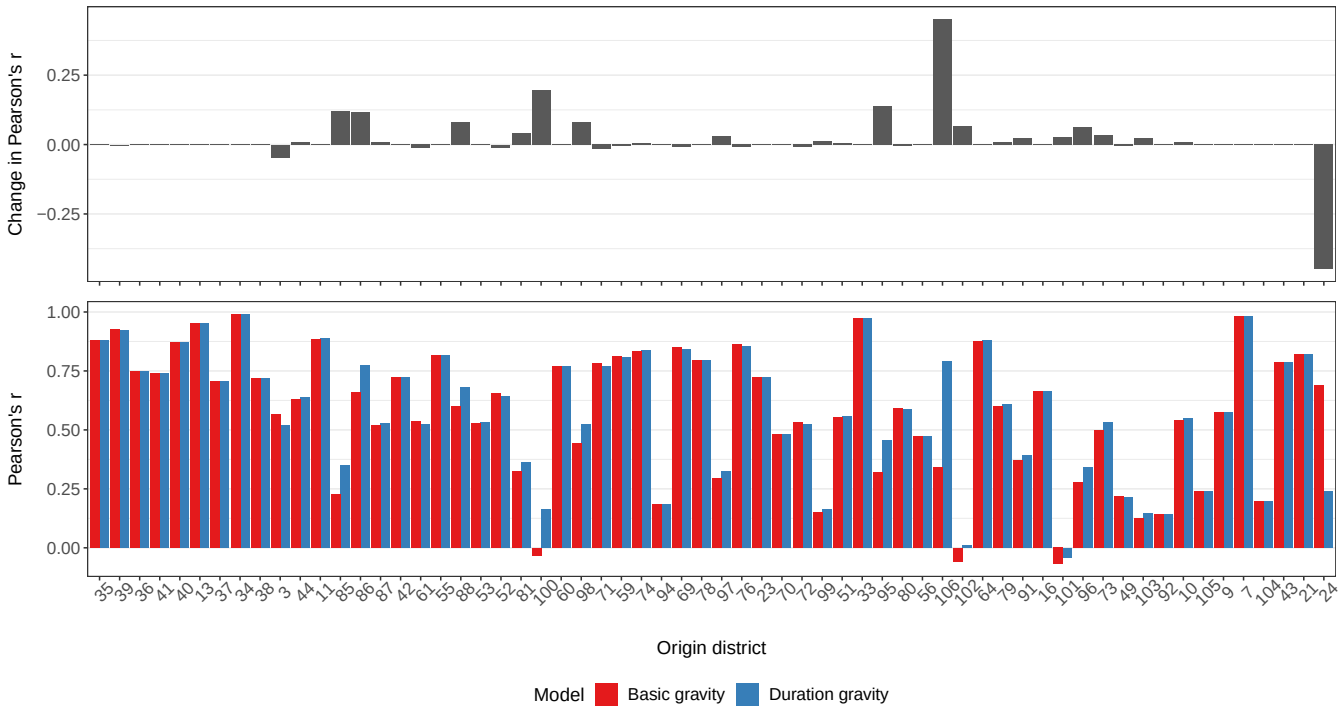

Fig. S10. Model performance of the trip duration decay model.

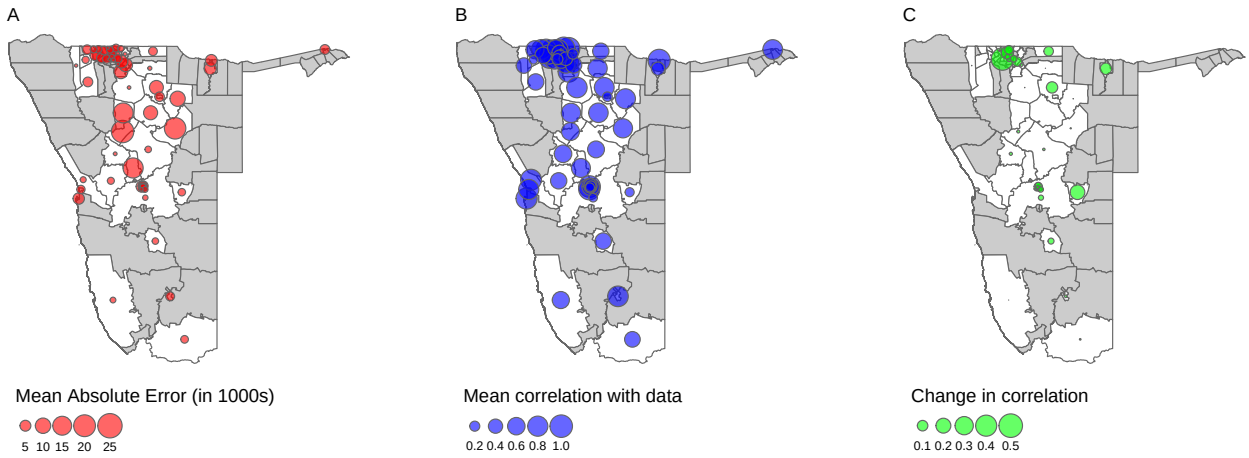

Fig. S11. Model performance of the trip duration decay model.

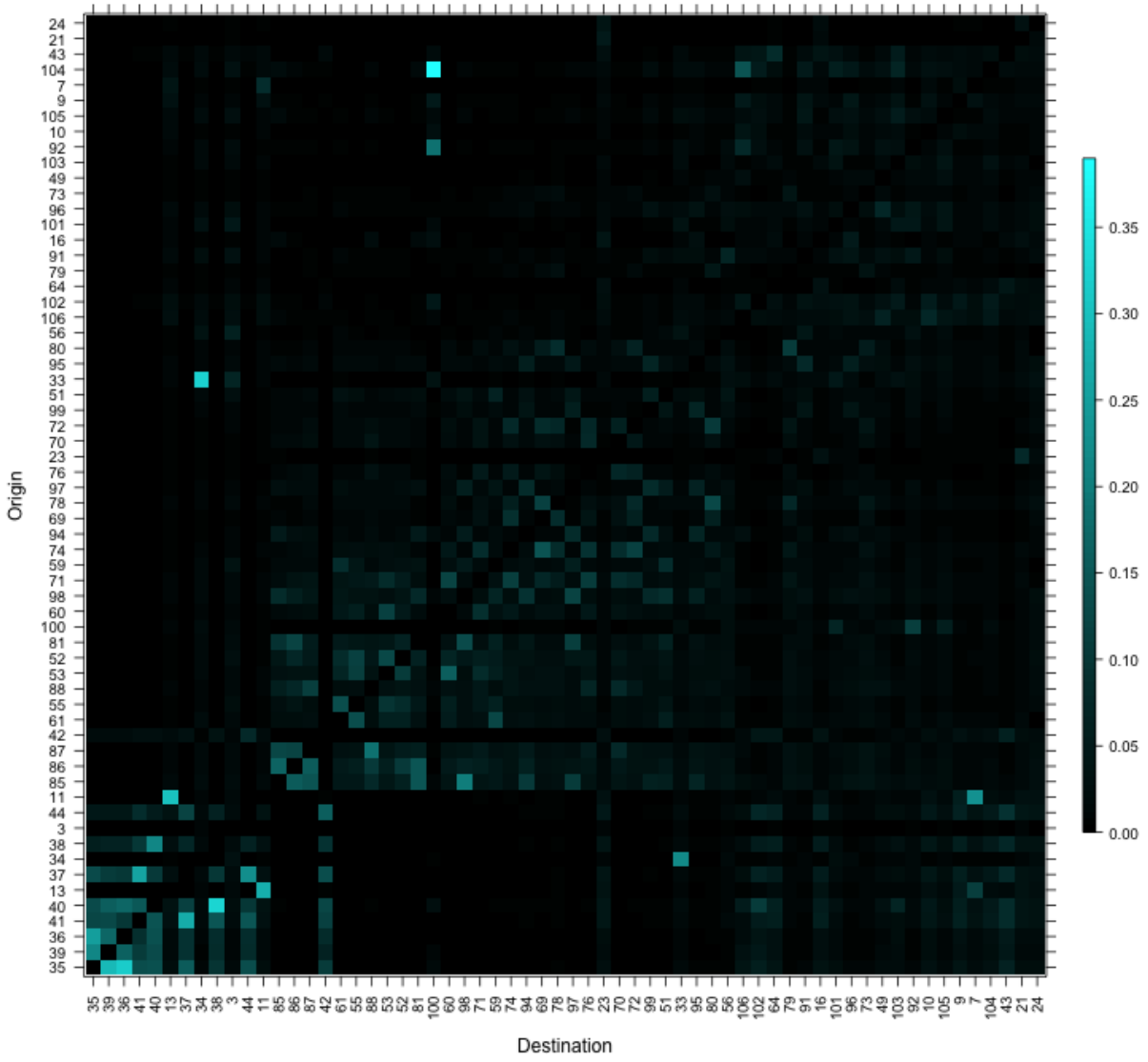

Fig. S12. Heatmap of pairwise values of estimated connectivity ($\hat{\pi}_{ij}$) among all 62 districts in the analysis. Districts are sorted from high population density to low density, where connectivity among high-density districts are in the lower left and connectivity among low-density districts is shown in the top right.

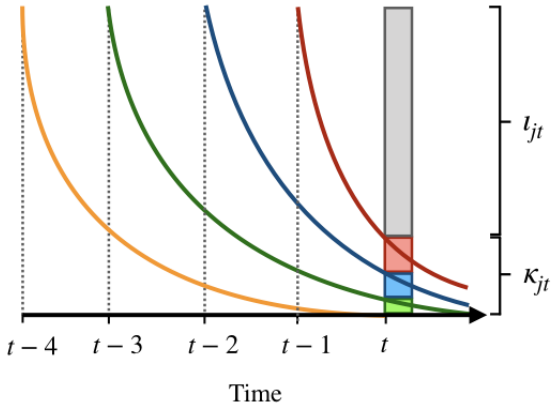

Fig. S13. A conceptual drawing of how the number of infected individuals from previous time steps (κ_{jt}) contribute to the spatial force of infection at time t of a simulation.

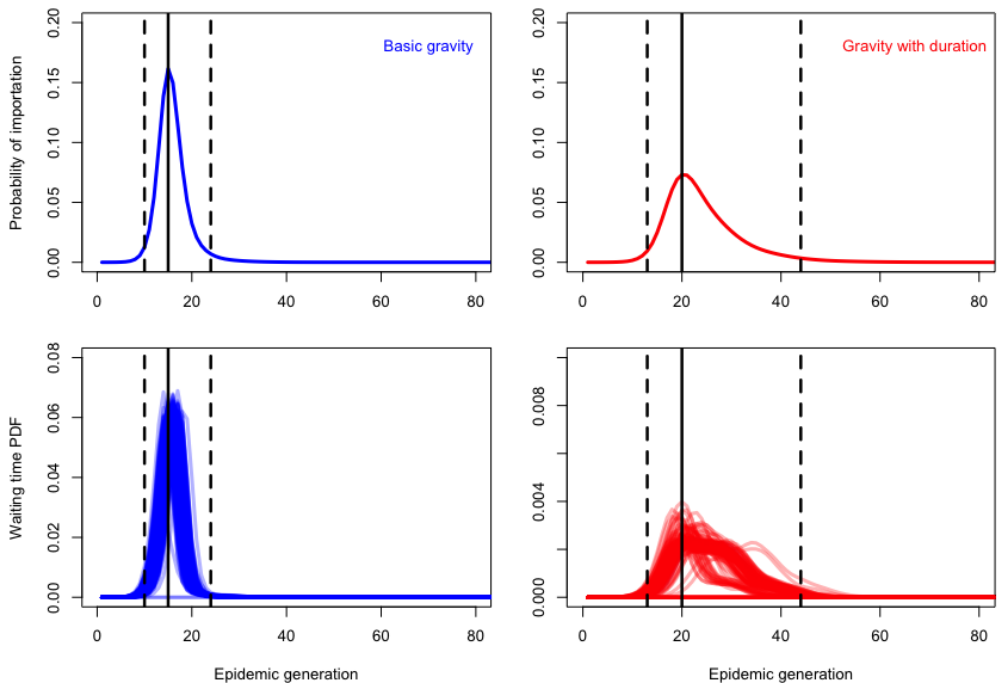

Fig. S14. An example of how simulations of the waiting time probability density function (PDF) are aggregated into the probability of pathogen importation. Spatial simulations using the basic gravity model are shown on the left (blue) and those with the gravity model with duration shown on the left (red). The solid lines indicate the peak and the dashed lines indicate the 95% highest posterior density (HPD) of the aggregate PDFs.

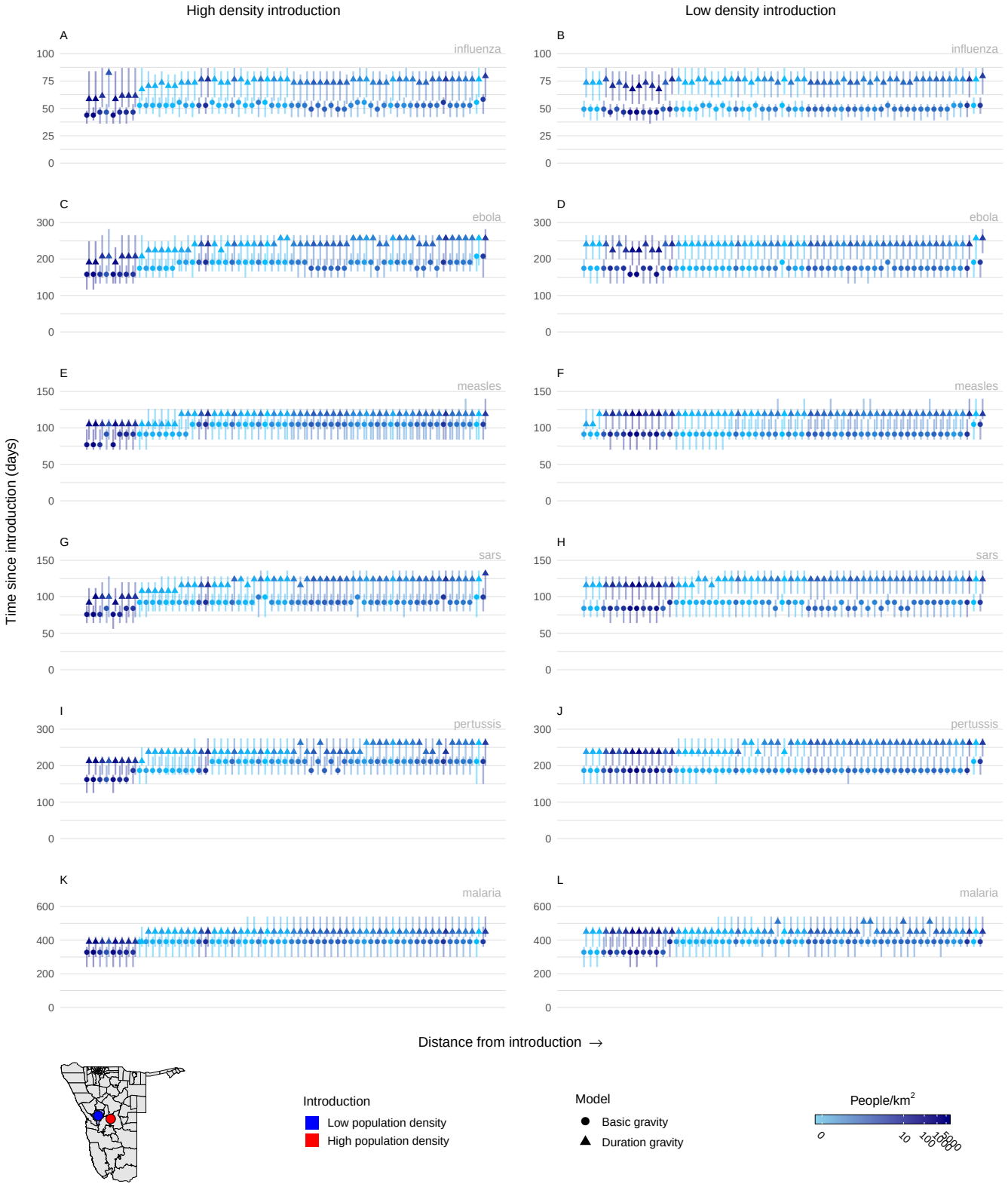

Fig. S15. Spatial TSIR simulations of infectious disease dispersal for 6 pathogens (influenza, measles, ebola, sars, pertussis, and malaria) introduced into districts with high population density (left) and low population density (right). Caterpillar plots represent the aggregated waiting time distributions of all simulations, where the peak waiting time is indicated with a circle for the basic gravity model or a triangle for the duration gravity model with vertical lines showing the 95% HPD intervals. The color of each caterpillar is given by the log population density of that district.

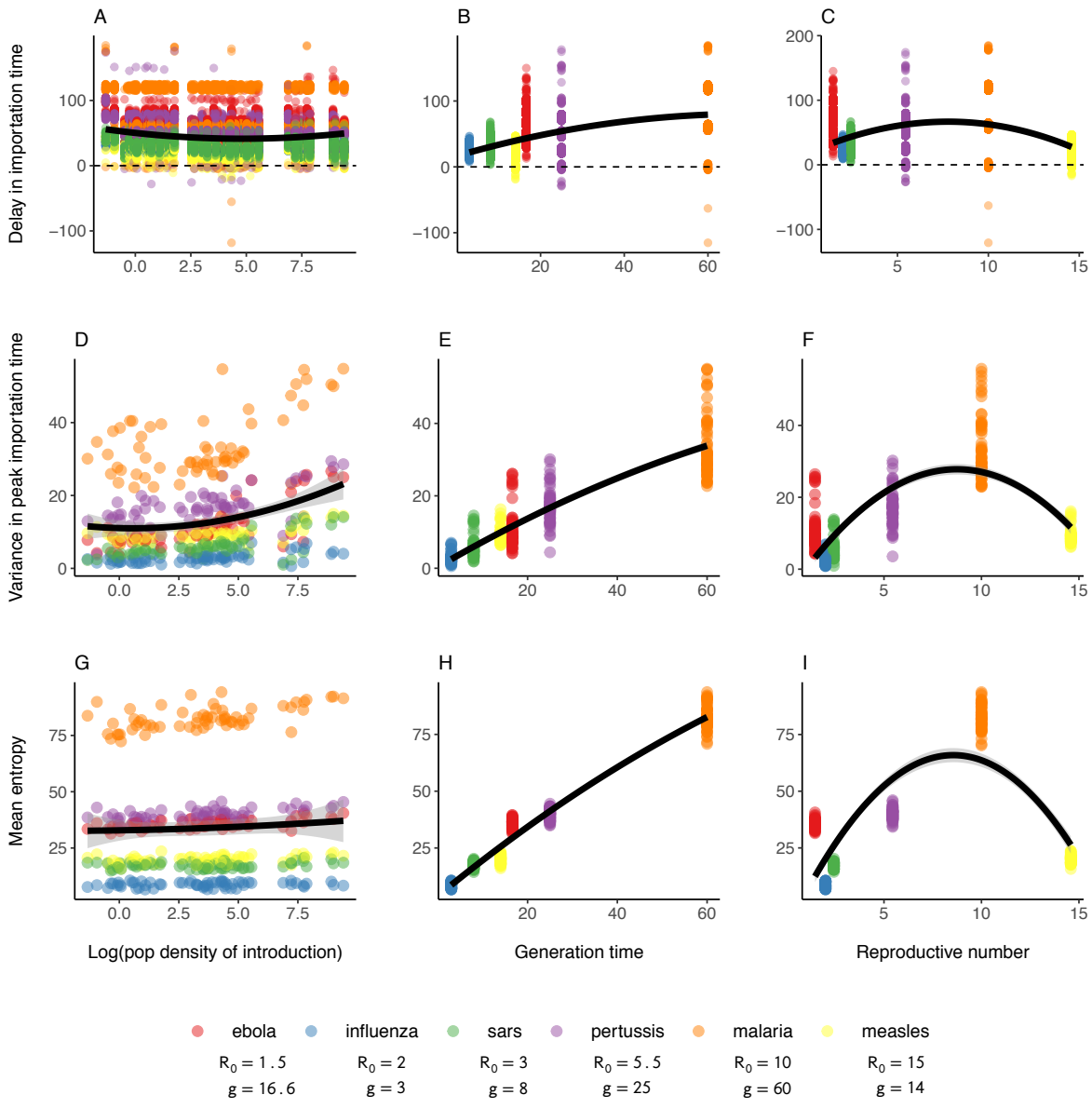

Fig. S16. Partial dependence plots showing overall patterns in spatial dispersal simulated by the duration TSIR model for 6 pathogens (influenza, measles, ebola, SARS, pertussis, and malaria). Results are plotted for three of the defining elements of simulations scenarios on the x-axes (population density of the introduction district, pathogen generation time g, and basic reproductive number R_0) and three metrics of spatial transmission dynamics on the y-axes (delay in importation times caused by incorporating trip duration, variance in peak importation time, and uncertainty in importation time measured by mean entropy of waiting time distributions). Pathogens are indicated by color and the overall trend among the plotted variables is shown with a LOESS trendline (black lines).

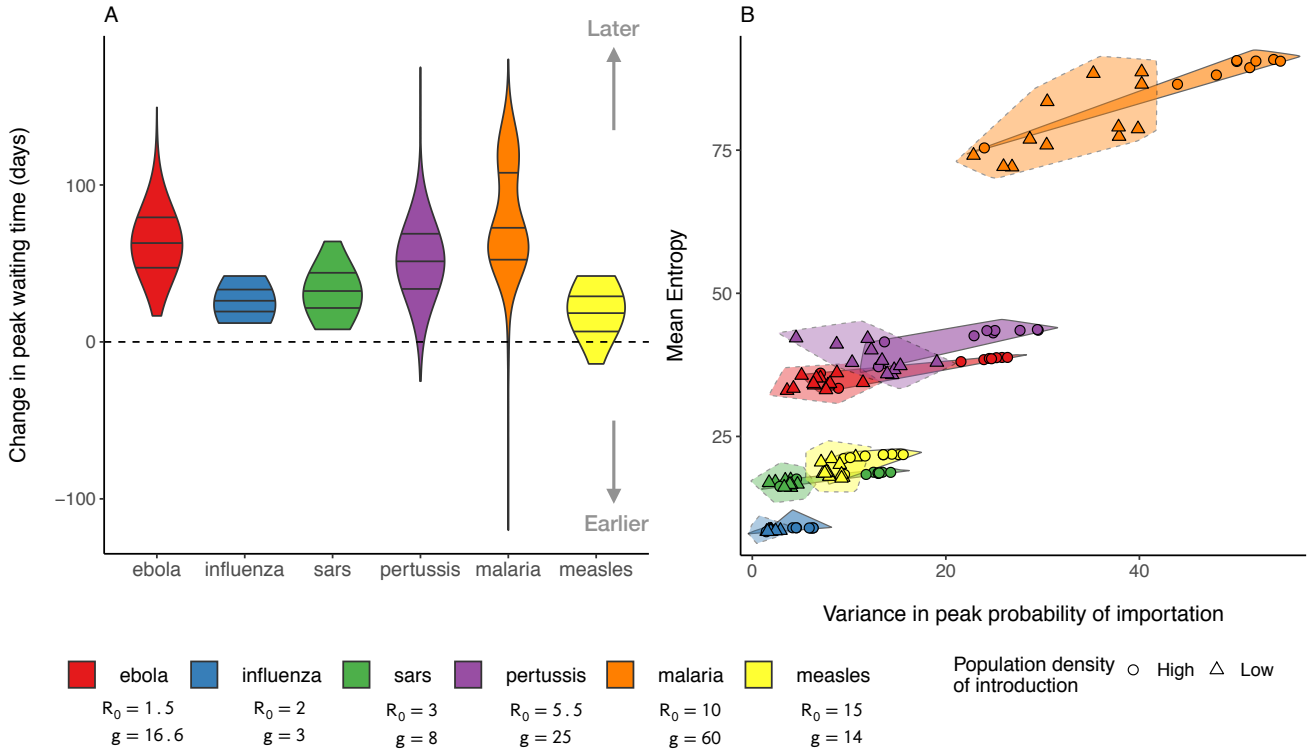

Fig. S17. Overall patterns in simulations of spatial transmission. Panel A shows the changes in peak waiting time distributions for 6 pathogens (influenza, measles, ebola, SARS, pertussis, and malaria) when the duration gravity model is used. Panel B shows the relationship between the variance and uncertainty in spatial spread given an introduction in the 10 highest (circles) and 10 lowest (triangles) density districts. Minimum convex polygons encircle each point type.

Table S1. Cumulative proportion of total trips for given values of trip duration (days) and trip distance (km).

Variable	Value	Proportion total trips
Duration (days)	1	0.47
	3	0.71
	7	0.85
	14	0.92
	30	0.96
	60	0.98
	90	0.99
Distance (km)	5	0.07
	10	0.16
	25	0.46
	50	0.63
	100	0.79
	250	0.92
	500	0.97

Table S2. Summary statistics for the distribution of estimated trip duration decay rate parameters ($\hat{\lambda}_{ij}$) for each route-type at two population density thresholds (1000 and 2500 people/km^2).

Threshold	Route-type	Mean $\hat{\lambda}_{ij}$	95% HPD
1000 people/km^2	High to high	0.45	0.22–0.53
	High to low	0.39	0.12–0.58
	Low to high	0.38	0.2–0.48
	Low to low	0.38	0.01–0.57
2500 people/km^2	High to high	0.48	0.45–0.53
	High to low	0.39	0.12–0.57
	Low to high	0.43	0.28–0.5
	Low to low	0.38	0.01–0.57

Table S3. Transmission parameters used in spatial TSIR simulations.

R_0	g	β	γ	Pathogen	Citation
10	60	30	3	malaria	(9, 10)
15	14	17.5	1.2	measles	
5.5	25	9.8	1.8	pertussis	(11, 12)
3	8	3.4	1.4	SARS	(13)
2	3	1.5	0.75	influenza	
1.5	16.6	3.8	2.6	Ebola	(14)

References

1. CJE Metcalf, et al., Implications of spatially heterogeneous vaccination coverage for the risk of congenital rubella syndrome in south africa. *J. Royal Soc. Interface* **10** (2013).
2. BT Grenfell, ON Bjørnstad, BF Finkenstädt, Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the tsir model. *Ecol. Monogr.* **72**, 185–202 (2002).
3. ON Bjornstad, BF Finkenstadt, BT Grenfell, Dynamics of measles epidemics: Estimating scaling of transmission rates using a time series SIR model. **72**, 17 (2002).
4. ON Bjørnstad, BT Grenfell, Hazards, spatial transmission and timing of outbreaks in epidemic metapopulations. *Environ. Ecol. Stat.* **15**, 265–277 (2008).
5. BF Finkenstädt, BT Grenfell, Time series modelling of childhood diseases: a dynamical systems approach. *J. Royal Stat. Soc. Ser. C Appl. Stat.* **49**, 187–205 (2000).
6. Y Xia, ON Bjørnstad, BT Grenfell, Measles metapopulation dynamics: A gravity model for epidemiological coupling and dynamics. *The Am. Nat.* **164**, 267–281 (2004).
7. K Glass, Y Xia, BT Grenfell, Interpreting time-series analyses for continuous-time biological models—measles as a case study. *J. Theor. Biol.* **223**, 19–25 (2003).
8. RT Clemen, RL Winkler, Combining probability distributions from experts in risk analysis. *Risk Analysis* **19** (1999).
9. JH Huber, GL Johnston, B Greenhouse, DL Smith, TA Perkins, Quantitative, model-based estimates of variability in the generation and serial intervals of plasmodium falciparum malaria. *Malar. J.* **15**, 490 (2016).
10. DL Smith, FE McKenzie, RW Snow, SI Hay, Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLOS Biol.* **5**, e42 (2007).
11. MA Vink, MCJ Bootsma, J Wallinga, Serial intervals of respiratory infectious diseases: A systematic review and analysis. *Am. J. Epidemiol.* **180**, 865–875 (2014).
12. DE te Beest, et al., Estimation of the serial interval of pertussis in dutch households. *Epidemics* **7**, 1–6 (2014).
13. RM Anderson, et al., Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic. *Philos. Transactions Royal Soc. London. Ser. B: Biol. Sci.* **359**, 1091–1105 (2004).
14. G Chowell, H Nishiura, Transmission dynamics and control of ebola virus disease (EVD): a review. *BMC Medicine* **12**, 196 (2014).