

A multi-season epidemic model with random genetic drift and transmissibility

Tom Britton^{1*} and Andrea Pugliese²

^{1*}Department of Mathematics, Stockholm University, Sweden.

²Department of Mathematics, University of Trento, Italy.

*Corresponding author(s). E-mail(s): tom.britton@math.su.se;

Contributing authors: andrea.pugliese@unitn.it;

Abstract

We consider a model for an influenza-like disease, in which, between seasons, the virus makes a random genetic drift δ , (reducing immunity by the factor δ) and obtains a new random transmissibility τ (closely related to R_0). Given the immunity status at the start of season k : $\mathbf{p}^{(k)}$, describing community distribution of years since last infection, and their associated immunity levels $\iota^{(k)}$, the outcome of the epidemic season k , characterized by the effective reproduction number $R_e^{(k)}$ and the fractions infected in the different immunity groups $\mathbf{z}^{(k)}$, is determined by the random pair (δ_k, τ_k) . It is shown that the immunity status $(\mathbf{p}^{(k)}, \iota^{(k)})$, is an ergodic Markov chain, which converges to a stationary distribution $\bar{\pi}(\cdot)$. More analytical progress is made for the case where immunity only lasts for one season. We then characterize the stationary distribution of $\mathbf{p}_1^{(k)}$, being identical to $\mathbf{z}^{(k-1)}$. Further, we also characterize the stationary distribution of $(R_e^{(k)}, \mathbf{z}^{(k)})$, and the conditional distribution of $\mathbf{z}^{(k)}$ given $R_e^{(k)}$. The effective reproduction number $R_e^{(k)}$ is closely related to the initial exponential growth rate $\rho^{(k)}$ of the outbreak, a quantity which can be estimated early in the epidemic season. As a consequence, this conditional distribution may be used for predicting the final size of the epidemic based on its initial growth and immunity status.

1 Introduction

Several infectious diseases are globally endemic but circulate around the globe leading to seasonal outbreaks in the winter seasons in the northern and southern hemisphere,

seasonal influenza being the prime example. Coming into a new epidemic season, a given community has some natural immunity originating from people being infected in recent years. Vaccinations, typically among the elderly population, reduce the number of severe cases but only marginally affect disease spreading, and are hence neglected in what follows. Between seasons, the circulating virus has often evolved (drifted), leading to decay of immunity among previously infected, and the transmissibility of the new strain may also differ from the transmissibility of the previous strain (disease severity may also differ, but this is not considered here).

One of the first to analyse mathematical models for these type of situations was Viggo Andreasen who wrote an influential paper where he analyses a deterministic model for seasonal outbreaks in which the immunity of individuals in the community decays in a specified deterministic way with the number of years since the last infection (Andreasen, 2003). When considering the special case where immunity only lasts one or two years he characterized the possible scenarios. For immunity lasting only one year he showed that there are three parameter regions in which, respectively, the yearly outbreak size converges to a constant, there is an outbreak only every other year, or the outbreak size alternates between a small and a large outbreak every other year. This model has later been extended in several ways. Kucharski et al (2016) consider a virus appearing in different variants, possibly having (partial) cross immunity to each other, and where immunity of the different strains depend on time since last infection of this strain. Roberts et al. (2024) extend the model by assuming a heterogeneous community (specified by a continuous distribution) affecting the immunity and infectivity of the individuals. Additionally they study what happens with the long term behaviour of the season outbreaks if the yearly deterministic immunity drift is perturbed randomly by a small amount each year.

In the current paper we take a more genuinely stochastic perspective by treating the yearly genetic drift, and also a transmissibility, (δ_k, τ_k) as a bivariate random vector, and where we assume the drift and transmissibility of different years to be independent. One reason for introducing such randomness is to allow for variable outcomes of seasonal outbreaks (as empirically observed), but also because the drift and transmissibility of a new strain depends on mutations in many loci on the virus, their common effect being highly complicated and thus conveniently modelled by random variables. The drift and transmissibility a given year may very well depend on each other, perhaps most likely having negative correlation (one of them being large may be enough to become dominant). Another novel feature in our model is that we allow also the transmissibility to vary between seasons. As a consequence, two quantities of the seasonal outbreaks will vary: the initial growth *and* the final epidemic size, as also observed empirically [3]. Considering a stochastic model also affect the analysis and in particular the resulting yearly epidemic outcome. Rather than converging to a fixed point or a deterministic periodic situation as in the deterministic setting, the stochastic model leads to a yearly epidemic outbreak depending on previous outbreaks but still being random, having some limiting stationary distribution depending on the random vector (δ, τ) .

In Section 2 we define the seasonal epidemic model with random genetic drift and transmissibility, and also characterize it as a Markov chain with a stationary distribution. In Section 3 we treat the case where immunity only lasts one season, for which more analytical progress is made. In Section 4 we present simulations for both when immunity lasts one or several years. The paper ends with a short discussion.

2 The multi-season model with random genetic drift and transmissibility

2.1 Model definition

The seasonal epidemic is driven by the yearly random genetic drift δ_k and the yearly random transmissibility τ_k , where the random pairs $(\delta_1, \tau_1), (\delta_2, \tau_2), \dots$ are assumed independent and identically distributed, but where δ_k and τ_k may depend on each other (in applications possibly being negatively correlated).

We assume that the immunity of an individual coming into season k depends on how many years j ago the individual was last infected, and that this immunity $\iota_j^{(k)}$ is given by the previous $j - 1$ genetic drifts as follows:

$$\iota_1^{(k)} = 1, \quad \iota_j^{(k)} = (1 - \delta_{k-1}) \cdots (1 - \delta_{k-j+1}), \quad 2 \leq j < r, \quad \text{and } \iota_r^{(k)} = 0,$$

meaning that immunity decays multiplicatively year by year, each year proportionally to the genetic drift. We make an additional assumption: that all immunity is lost after $r \geq 2$ years, so $\iota_r^{(k)} = 0$ for all k irrespective the genetic drifts.

What happens during outbreak season k depends on the drift and transmissibility (δ_k, τ_k) but also on the community distribution of when individuals were last infected: we introduce $\mathbf{p}^{(k)} = (p_1^{(k)}, \dots, p_r^{(k)})$, where $p_j^{(k)}$ denotes the community fraction which were last infected j seasons ago, and $p_r(k) = 1 - \sum_{j=1}^{r-1} p_j^{(k)}$ includes also those who were infected more than r seasons ago, or never: all of them lacking any immunity. Given the vector $\mathbf{p}^{(k)}$ and their associated immunity levels $\boldsymbol{\iota}^{(k)} = (\iota_1^{(k)}, \dots, \iota_{r-1}^{(k)})$ and the drift and transmissibility of year k : (δ_k, τ_k) , the effective reproduction number is given by

$$R_E^{(k)} = \tau_k \sum_{j=1}^r p_j^{(k)} \left(1 - (1 - \delta_k) \iota_j^{(k)} \right). \quad (1)$$

The basic reproduction number, corresponding to the situation where there is no immunity, is simply $R_0^{(k)} = \tau_k$. Further, the final epidemic outcome season k in the different immunity classes, $\mathbf{z}^{(k)} = (z_1^{(k)}, \dots, z_r^{(k)})$, is determined by the drift and transmissibility year k : (δ_k, τ_k) through the following final size equations

$$\begin{aligned} 1 - z_1 &= e^{-(1 - (1 - \delta_k)) \tau_k (\sum_j p_j^{(k)} z_j)} \\ &\dots \\ 1 - z_j &= e^{-(1 - (1 - \delta_k) \iota_j^{(k)}) \tau_k (\sum_j p_j^{(k)} z_j)} \end{aligned}$$

$$\dots$$

$$1 - z_r = e^{-\tau_k (\sum_j p_j^{(k)} z_j)},$$

It is well known that these equations have a unique positive solution $\mathbf{z}^{(k)}$ iff $R_E^{(k)} > 1$ (e.g. Sec 13.2.3 in [1]). The component $z_j^{(k)}$ denotes the fraction infected during epidemic season k among those who were last infected j seasons ago. The overall fraction getting infected is hence $z^{(k)} = \sum_j z_j^{(k)} p_j^{(k)}$. Summing the equations above, weighted by $p_j^{(k)}$, gives the following equation for the overall fraction getting infected

$$1 - z = \sum_j p_j^{(k)} e^{-(1-(1-\delta_k)\iota_j^{(k)})\tau_k z}. \quad (2)$$

This is a 1-dimensional equation for $z^{(k)}$, the overall fraction getting infected, from which the fraction infected of the different types is easily obtained: $z_j^{(k)} = 1 - e^{-(1-(1-\delta_k)\iota_j^{(k)})\tau_k z^{(k)}}$.

2.2 Characterization as a Markov chain and its stationary distribution

We now describe the epidemic model as a Markov chain. Given the integer $r \geq 2$ indicating how many years until immunity is completely lost, the state of our Markov chain is given by the vector $(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})$, where $\mathbf{p}^{(k)} = (p_1^{(k)}, \dots, p_r^{(k)})$ describes the community distribution of time since last infection, and $\boldsymbol{\iota}^{(k)}$ their associated immunity levels $\boldsymbol{\iota}^{(k)} = (\iota_1^{(k)}, \dots, \iota_{r-1}^{(k)})$, with $\iota_j^{(k)} = (1 - \delta_{k-1}) \cdots (1 - \delta_{k-(j-1)})$. The vector $(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})$ in fact only has dimension $2r - 3$, since $p_r^{(k)} = 1 - \sum_{j=1}^{r-1} p_j^{(k)}$ and $\iota_1^{(k)} = 1$. We now show that the vector $\{(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})\}_{k=1}^\infty$ is a Markov chain driven by (δ_k, τ_k) . Given $(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})$, the next state in the Markov chain, $(\mathbf{p}^{(k+1)}, \boldsymbol{\iota}^{(k+1)})$, is determined by (δ_k, τ_k) as follows. The final outcome of the epidemic season k : $\mathbf{z}^{(k)} = (z_1^{(k)}, \dots, z_r^{(k)})$ is determined as described in the previous subsection (if $R_E^{(k)} \leq 1$ then all $z_j^{(k)} = 0$ and otherwise they are all strictly positive). The fraction having j years since last infection coming into the next epidemic season $k + 1$ is then equal to the fraction having $j - 1$ years since last infection in season k and who avoid getting infected in season k : $p_j^{(k+1)} = p_{j-1}^{(k)} (1 - z_{j-1}^{(k)})$. The first and last components are different: $p_r^{(k+1)} = p_r^{(k)} (1 - z_r^{(k)}) + p_{r-1}^{(k)} (1 - z_{r-1}^{(k)})$, and $p_1^{(k+1)} = \sum_{j=1}^r p_j^{(k)} z_j^{(k)}$ (=all who got infected season k). The immunity levels are simply shifted one step: $\iota_j^{(k+1)} = \iota_{j-1}^{(k)} (1 - \delta_k)$, and $\iota_1^{(k+1)} = 1$ and $\iota_r^{(k)} = 0$.

We have hence shown that $(\mathbf{p}^{(k+1)}, \boldsymbol{\iota}^{(k+1)})$ is determined by (δ_k, τ_k) given $(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})$, implying that $\{(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})\}_{k=1}^\infty$ is a Markov chain. A Markov chain is often described by its transition kernel $\pi(\cdot|\cdot)$. It is in principle possible to determine the transition matrix $\pi((\mathbf{p}^{(k+1)}, \boldsymbol{\iota}^{(k+1)})|(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)}))$, at least numerically, although it is quite complicated for

general r . In Section 3 we make more analytical progress for the case $r = 2$. Instead we end this section with the following theorem for our Markov chain, valid for arbitrary r .

Proposition 2.1. *Assume the random vector (δ, τ) has a continuous density function with strictly positive support on $[0, 1) \times [0, \infty)$, and with an atom at $\delta = 1$ (no immunity) with positive probability. Then the Markov chain $\{(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})\}_{k=1}^{\infty}$ is recurrent, aperiodic, and converges to a stationary distribution $\bar{\pi}(\cdot)$.*

Proof. We first prove recurrence. We do this by showing that any state $(\mathbf{p}, \boldsymbol{\iota})$ can reach the state $(0, \dots, 1, 1, 0, \dots, 0)$ with $p_r = 1$ and $\iota_1 = 1$ and all other components being 0 (so no immunity). This will happen if the next r random pairs (δ_j, τ_j) have $\delta_j = 1$ and $\tau_j < 1$, an event with positive probability ($\tau = R_0 < 1$ imply no outbreak).

From state $(0, \dots, 1, 1, 0, \dots, 0)$ we now show that any state $(\mathbf{p}, \boldsymbol{\iota})$ can be reached after r seasons. We do this by induction in r . Start with $r = 2$. We then want to find a τ such that the final outcome $z = p_1$ for some given value p_1 , starting with no prior immunity. This is easy: we simply choose τ_1 such that $1 - p_1 = e^{-p_1 \tau_1}$, i.e. $\tau_1 = -\ln(1 - p_1)/p_1$.

For the induction step we want to show that we can reach any fixed state $(p_1, \dots, p_r, \iota_1 = 1, \dots, \iota_{r-1})$ after r seasons, starting with no immunity season 0. We use that, by induction, we can reach any state $(p'_1, \dots, p'_{r-1}, \iota'_1 = 1, \dots, \iota'_{r-2})$ after $r - 1$ seasons. Our task is hence to specify a vector $(p'_1, \dots, p'_{r-1}, \iota'_1 = 1, \dots, \iota'_{r-2})$ and a drift and transmissibility (δ, τ) which result in the desired $(p_1, \dots, p_r, \iota_1 = 1, \dots, \iota_{r-1})$.

We start with the immunities. Set $\delta = 1 - \iota_2$, and $\iota'_j = \iota_{j+1}/(1 - \delta)$, $j = 1, \dots, r - 2$, leading to the desired immunities.

We now want to specify the vector (p'_1, \dots, p'_{r-1}) and transmissibility τ which, together with immunity vector $(\iota'_1 = 1, \dots, \iota'_{r-2})$ result in the vector (p_1, \dots, p_r) next year.

If $p_1 = 0$, implying that there was no outbreak the previous year, this is easy: we simply set $p'_j = p_{j+1}$, $j = 1, \dots, r - 1$, and we choose any $\tau < 1$ so that no outbreak will occur.

The case $p_1 > 0$ is a bit more involved. The vector (p_1, \dots, p_r) will be obtained by iteratively choosing τ_i , $i = 1, 2, \dots$ making our vector $(p_1^{(i)}, \dots, p_r^{(i)})$ approaching (p_1, \dots, p_r) closer and closer.

Start with some arbitrary value larger than 1, e.g. $\tau_1 = 1.5$.

The following step is repeated until suitable precision is obtained. For any i , set

$$1 - z'_j = e^{-(1 - (1 - \delta)\iota'_j)\tau_i p_1}, \text{ and } p'_{j-1} = p_j / (1 - z'_{j-1}), j = 1, \dots, r.$$

These candidate values for $\{z'_j\}$ (which depend on i through τ_i !) give the right relations between them (see end of the previous subsection), but it may give too few or too many overall infected. This we now control by Equation (2). Compute $1 - z' = \sum_j p'_j e^{-(1 - \iota'_j)\tau_i p_1}$. If $1 - p_1 > 1 - z'$ it means too many got infected overall for our candidate value τ_i , implying that we should pick a smaller τ_{i+1} and repeat the procedure. On the other hand, if $1 - p_1 < 1 - z'$ it means too few got infected overall for our candidate value τ_i , implying that we should pick a larger τ_{i+1} and repeat the procedure. In each iteration we can pick a new τ_i being half way between the previous τ_{i-1} and the closest earlier tested value in the desired direction, implying that our sequence τ_i will converge, and we stop when our vector (p'_1, \dots, p'_{r-1}) and τ_i comes close enough to the desired outcome (p_1, \dots, p_r) . This completes the induction step. The aperiodicity follows from the fact that the immunity-free state $(0, \dots, 1, 1, 0, \dots, 0)$ can return to itself with positive probability (if $\tau < 1$).

Since our Markov chain $\{(\mathbf{p}^{(k)}, \boldsymbol{\iota}^{(k)})\}_{k=1}^{\infty}$ returns to the immunity-free state $(0, \dots, 1, 1, 0, \dots, 0)$ in r steps with positive probability it is a Harris chain (cf. Sec 5.8

in Durrett, 2019). It then follows that it has stationary measure $\bar{\pi}(\cdot)$ to which the chain converges irrespective of starting value. \square

3 Analytical results for the case $r = 2$ where immunity is lost after two seasons

We now focus on the case where $r = 2$, so immunity only lasts the year after being infected. In this case our vector $(\mathbf{p}^{(k)}, \mathbf{t}^{(k)})$ is simply the one dimensional quantity $p_1^{(k)} =: p^{(k)}$, since the outcome on year k only depends on (δ_k, τ_k) and how many that were infected last season $p_1^{(k)} = z^{(k-1)}$. We will hence study the distribution of the outcome some year given the outcome of the previous year. We drop the k -notation and hence consider the outcome of R_E and z , given p . All results below are hence conditional on p , the fraction infected in the previous season.

We assume that joint probability distribution of the pair (δ, τ) has density q (with support in $(0, 1) \times (0, +\infty)$), and we wish to explicitly compute some properties of the Markov chain in this specific case.

3.1 Bivariate distribution of effective reproduction ratio and attack ratio

Two important properties of a seasonal epidemic are its effective reproduction ratio R_e and its final attack ratio z . The value of R_e determines the initial exponential growth rate, and can be estimated in the first phase of an epidemic. The final attack ratio z (i.e. the fraction of the population that gets infected in one season) determines (together with the infection severity) the impact of an epidemic; by definition, it can be ascertained only after an epidemic is over, and may be difficult to estimate, because of under-reporting and other concurring infections.

As stated above, when $r = 2$, the Markov chain is one-dimensional, and only depends on the fraction p infected in the previous year. Simplifying for the case $r = 2$ the formulae (1) and (2), we have

$$R_e = \tau(p\delta + 1 - p). \quad (3)$$

If $R_e \leq 1$, then $z = 0$ is the only solution, but if $R_e > 1$, there is a unique solution z in $(0, 1)$ to the equation

$$1 - z = pe^{-\tau\delta z} + (1 - p)e^{-\tau z}. \quad (4)$$

In order to consider the probability distribution of the pair (z, R_e) , given p (the fraction infected in the previous year), we need to distinguish the case $p = 0$ from $p > 0$. We obtain the following

Proposition 3.1. *Assume that the pair (δ, τ) has density $q(\cdot, \cdot)$ (with support in $(0, 1) \times (0, +\infty)$).*

If $p^{(k)} = 0$, the distribution of the pair $(z^{(k)}, R_e^{(k)})$ lies on the one-dimensional curve

$(z(R_e), R_e)$ where, if $R_e > 1$, $z(R_e)$ is the unique solution in $(0, 1)$ of $1 - z = e^{-R_e z}$, while $z(R_e) = 0$ if $R_e \leq 1$.

On that curve $R_e^{(k)}$ has density $q_R(y) = \int_0^1 q(x, y) dx$.

If $p^{(k)} = p > 0$, the distribution of $(z^{(k)}, R_e^{(k)})$ is composed of two parts. With probability

$$\mathbb{P}(z = 0|p) = \mathbb{P}(R_e \leq 1|p) = \int_0^1 \int_0^{1/(p\delta+1-p)} q(\delta, \tau) d\tau d\delta. \quad (5)$$

the pair lies in the segment $\{z = 0, 0 < R_e \leq 1\}$ with density

$$\int_0^1 q\left(\delta, \frac{R_e}{p\delta+1-p}\right) \frac{1}{p\delta+1-p} d\delta.$$

The second part lies in $\{0 < z < 1, R_e > 1\}$ with bivariate density

$$f(z, R_e|p) = \begin{cases} q(\delta^*, \tau^*) \frac{p\delta^*+1-p-R_e(p\delta^*e^{-\delta^*\tau^*z}+(1-p)e^{-\tau^*z})}{p(1-p)R_ez(e^{-\delta^*\tau^*z}-e^{-\tau^*z})} & \text{if } (z, R_e) \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where

$$\begin{aligned} \mathcal{A} = \{0 < z < 1-p, -\frac{\log(1-z)}{z} < R_e < -\frac{(1-p)}{z} \log\left(1 - \frac{z}{1-p}\right)\} \\ \cup \{1-p \leq z < 1, -\frac{\log(1-z)}{z} < R_e\}, \end{aligned} \quad (7)$$

$\delta^*(z, R_e)$ is the unique solution (existing when $(z, R_e) \in \mathcal{A}$) in $(0, 1)$ of

$$G(\delta, z, R_e) := pe^{-\frac{\delta R_e z}{p\delta+1-p}} + (1-p)e^{-\frac{R_e z}{p\delta+1-p}} + z - 1 = 0. \quad (8)$$

and

$$\tau^* = \frac{R_e}{p\delta^*+1-p}.$$

Proof If $p = 0$, formulae (3) and (4) show that $R_e = R_0 = \tau$ and z is the only solution of $1 - z = e^{-\tau z}$, which is completely determined by τ . Hence the pair (z, R_e) lies on a one-dimensional curve: R_e has density $q_R(y) = \int_0^1 q(x, y) dx$, while z is a function of R_e . The density of R_e is the density of τ which is obtained by marginalizing the density q as in the statement of the Theorem.

On the other hand, if $p > 0$, we can consider the map

$$\begin{aligned} F : (0, 1) \times (0, +\infty) &\rightarrow [0, 1) \times (0, +\infty) \\ (\delta, \tau) &\rightarrow (z, R_e). \end{aligned}$$

Using the standard formula (see. for instance, [5, Theorem 12.7]) for the density of a random variable $F(X)$, the density of the pair (z, R_e) is

$$f(z, R_e) = \begin{cases} q(F^{-1}(z, R_e)) \times |\det((F^{-1})'(z, R_e))| & \text{if } (z, R_e) \in \text{Im}(F) \\ 0 & \text{if } (z, R_e) \notin \text{Im}(F). \end{cases} \quad (9)$$

In this case, $F_1((\delta, \tau)) \equiv 0$ for all values such that $F_2((\delta, \tau)) \leq 1$; thus F is not invertible everywhere. It is useful to think of (9) as a defective density restricted to the set $\{(z > 0, R_e > 1)\}$, while, when $R_e \leq 1$ the support of $(z^{(k+1)}, R_e^{(k+1)})$ lies in the segment with $z = 0$. It is easy to see that $\mathbb{P}(z = 0) = \mathbb{P}(R_e \leq 1)$ can be computed as in (5), while on that segment the density of $R_e^{(k+1)}$ can be computed from

$$\mathbb{P}(R_e^{(k+1)} \leq R_e) = \int_0^1 \int_0^{R_e/(p\delta+1-p)} q(\delta, \tau) d\tau d\delta.$$

Considering now (9), the density $f(z, R_e)$ cannot be written explicitly, since there are not explicit expressions for F and F^{-1} . However, one can improve on (9) by computing the range of F and the derivatives of F^{-1} .

In order to find the range of F , we first find

$$\tau = \frac{R_e}{p\delta + 1 - p}$$

and then, substituting in (4), δ can be obtained as a solution of $G(\delta, z, R_e) = 0$ with G shown in (8).

Simple computations then yield

$$\frac{\partial G}{\partial \delta} = -\frac{p(1-p)R_e z}{(p\delta + 1 - p)^2} \left(e^{-\frac{\delta R_e z}{p\delta + 1 - p}} - e^{-\frac{R_e z}{p\delta + 1 - p}} \right) < 0 \quad (10)$$

$$\frac{\partial G}{\partial R_e} = -\frac{z}{p\delta + 1 - p} \left(p\delta e^{-\frac{\delta R_e z}{p\delta + 1 - p}} + (1-p)e^{-\frac{R_e z}{p\delta + 1 - p}} \right) < 0 \quad (11)$$

$$\frac{\partial G}{\partial z} = 1 - \frac{R_e}{p\delta + 1 - p} \left(p\delta e^{-\frac{\delta R_e z}{p\delta + 1 - p}} + (1-p)e^{-\frac{R_e z}{p\delta + 1 - p}} \right) > 0. \quad (12)$$

The sign of $\frac{\partial G}{\partial z}$ comes from the fact that the derivative of the RHS of (4) is larger than -1 at the solution $z \in (0, 1)$ of (4).

Since G is decreasing a function of δ , there exists a solution $\delta^*(z, R_e)$ of $G(\delta, z, R_e) = 0$ if and only if

$$p + (1-p)e^{-\frac{R_e z}{1-p}} + z - 1 = G(0, z, R_e) > 0 > G(1, z, R_e) = e^{-R_e z} + z - 1. \quad (13)$$

It is easy to see that, if $z < 1 - p$, (13) is equivalent to

$$-\frac{\log(1-z)}{z} < R_e < -\frac{(1-p)}{z} \log \left(1 - \frac{z}{1-p} \right),$$

while only the first condition holds if $z \geq 1 - p$.

It follows that

$$\text{Im}(F) = \mathcal{A} \cup \{z = 0, R_e \leq 1\},$$

with \mathcal{A} specified in (7).

Considering only the case where $R_e > 1$, where F is invertible and (9) holds, we have

$$F^{-1}(z, R_e) = \left(\delta^*(z, R_e), \frac{R_e}{p\delta^*(z, R_e) + 1 - p} \right).$$

We can then compute its Jacobian using (10)–(12). First of all, notice that

$$\frac{\partial \tau}{\partial z} = -\frac{pR_e}{(p\delta^* + 1 - p)^2} \frac{\partial \delta}{\partial z} \quad \frac{\partial \tau}{\partial R_e} = \frac{1}{p\delta^* + 1 - p} - \frac{pR_e}{(p\delta^* + 1 - p)^2} \frac{\partial \delta}{\partial R_e}.$$

Then

$$\begin{vmatrix} \frac{\partial \delta}{\partial z} & \frac{\partial \delta}{\partial R_e} \\ \frac{\partial \tau}{\partial z} & \frac{\partial \tau}{\partial R_e} \end{vmatrix} = \frac{\partial \delta}{\partial z} \cdot \frac{1}{p\delta^* + 1 - p}.$$

Using the implicit function theorem, and remembering $\tau = \frac{R_e}{p\delta + 1 - p}$,

$$\frac{\partial \delta}{\partial z} = -\frac{\frac{\partial G}{\partial z}}{\frac{\partial G}{\partial \delta}} = (p\delta^* + 1 - p) \frac{p\delta^* + 1 - p - R_e \left(p\delta^* e^{-\delta^* \tau^* z} + (1-p)e^{-\tau^* z} \right)}{p(1-p)R_e z (e^{-\delta^* \tau^* z} - e^{-\tau^* z})}.$$

Substituting, we obtain (6). □

3.2 Transition probabilities

From (6) one can compute several other formulae of interest. We start with the transition probabilities. Under the assumption that (δ, τ) has density q (with support in $(0, 1) \times (0, +\infty)$), one sees that $p^{(k)}$ has a distribution with an atom at 0 and a density over $(0, 1)$, and the same will be true for $p^{(k+1)}$. Then one obtains the following

Proposition 3.2. *The transition probability of the Markov chain consist of*

$$\mathbb{P}(p^{(k+1)} = 0 | p^{(k)} = 0) = \pi_{00} = \mathbb{P}(\tau \leq 1) = \int_0^1 \int_0^1 q(x, y) dx dy. \quad (14)$$

$$\begin{aligned} \mathbb{P}(p^{(k+1)} = 0 | p^{(k)} = p) &= \pi(p, 0) = \mathbb{P}(R_e^{(k)} \leq 1) = \mathbb{P}(\tau(\delta p + (1-p)) \leq 1) \\ &= \int_0^1 \int_0^{1/(xp + (1-p))} q(x, y) dy dx. \end{aligned} \quad (15)$$

while the density function of $p^{(k+1)}$ for $p' > 0$ conditional on $p^{(k)} = p \geq 0$ is

$$\pi(p, p') = \int_0^1 q(x, \tau^*) \frac{1 - \tau^* \left(p x e^{-\tau^* x p'} + (1-p)e^{-\tau^* p'} \right)}{p' (p x e^{-\tau^* x p'} + (1-p)e^{-\tau^* p'})} dx \quad (16)$$

where $\tau^* = \tau^*(x, p')$ is the unique solution of

$$p e^{-x \tau p'} + (1-p)e^{-\tau p'} + p' - 1 = 0.$$

Proof The computation of (14) and (15) is straightforward.

To compute the density function $\pi(p, p')$ where $p' \in (0, 1)$ and $p \geq 0$, one needs only integrate (6) to obtain

$$\pi(p, p') = \begin{cases} -\frac{(1-p)}{p'} \log\left(1 - \frac{p'}{1-p}\right) \int_{-\frac{\log(1-p')}{p'}}^{\frac{(1-p)}{p'} \log\left(1 - \frac{p'}{1-p}\right)} f(p', y) dy & p' < 1-p \\ \int_{-\frac{\log(1-p')}{p'}}^{+\infty} f(p', y) dy & p' \geq 1-p. \end{cases} \quad (17)$$

It is more convenient to change the integration variable in (17) to $x = \delta^*(p', y)$; then

$$dy = -\frac{\frac{\partial G}{\partial \delta}}{\frac{\partial G}{\partial R_e}} dx$$

while

$$f(p', y) = -q(\delta^*(p', y), \tau^*(p', y)) \frac{\frac{\partial G}{\partial z}}{\frac{\partial G}{\partial \delta}} \cdot \frac{1}{p\delta^* + 1 - p}. \quad (18)$$

Hence, using (10)–(12),

$$\begin{aligned} f(p', y) dy &= -q(x, \tau^*) \frac{\frac{\partial G}{\partial z}}{\frac{\partial G}{\partial R_e}} \cdot \frac{1}{px + 1 - p} dx \\ &= q(x, \tau^*) \frac{1 - \tau^* (pxe^{-\tau^* x p'} + (1-p)e^{-\tau^* p'})}{p' (pxe^{-\tau^* x p'} + (1-p)e^{-\tau^* p'})} dx \end{aligned}$$

Substituting in (17), we obtain (16). \square

3.3 The conditional distribution of final size given the effective reproduction ratio

As mentioned above, coming into season k the fraction infected in the previous season $z^{(k-1)} = p^{(k)}$ is known, and the value of $R_e^{(k)}$ can be estimated from data during the early phase of the epidemic, while the final size $z^{(k)}$ ($= p^{(k+1)}$) can be estimated from data only at the end of a seasonal epidemic. Hence it may be useful to predict $z^{(k)}$ knowing $z^{(k-1)} = p^{(k)}$ and $R_e^{(k)}$.

If $R_e^{(k)} \leq 1$, necessarily $z^{(k)} = 0$. Similarly if $z^{(k-1)} = 0$, there is no immunity; then both $R_e^{(k)}$ and $z^{(k)}$ are determined by τ_k only, so that knowing one, the other one can be exactly computed.

Finally, the non-trivial case is where $z^{(k-1)} > 0$ and $R_e^{(k)} = R_e > 1$. We wish to compute the conditional density

$$f(z^{(k)} = z | R_e^{(k)} = R_e, p^{(k-1)} = p) = \frac{f(z, R_e | p)}{\int_{z': (z', R_e) \in \mathcal{A}} f(z', R_e | p) dz'}, \quad (19)$$

where f denotes the density for whatever variables and conditional events that are indicated.

In this case too, it is convenient to change the integration variable to $x = \delta^*(p', R_e)$ so that $dp' = -\frac{\frac{\partial G}{\partial \delta}}{\frac{\partial G}{\partial z}} dx$. Remembering that δ^* is the solution of $G(\delta, z, R_e) = 0$ as defined in (8), this yields, together with (18),

$$f(z|R_e, p) = \frac{f(z, R_e|p)}{\int_0^1 q(x, \tau) \frac{1}{px+1-p} dx} \quad (20)$$

where in the integral $\tau = \frac{R_e}{px+1-p}$.

3.4 Stationary distribution

The results of Section 2 can be made more specific when $r = 2$. Precisely

Proposition 3.3. *Assume that the pair (δ, τ) has a density q with support in $(0, 1) \times (\tau_m, \tau_M)$ with $0 \leq \tau_m < 1 < \tau_M \leq +\infty$, such that the marginal density $q_\tau(y) = \int_0^1 q(x, y) dx > 0$ on (τ_m, τ_M) .*

Then the Markov chain $\{p^{(k)}\}_{k=1}^\infty$ is recurrent, aperiodic, and converges to a stationary distribution $\bar{\pi}$ which consists of an atom $\bar{\pi}_0$ at 0 and a density $\bar{\pi}(\cdot)$ with support on $(0, \bar{z})$; $\bar{z} = 1$ if $\tau_M = +\infty$, while \bar{z} is the solution of $1 - z = e^{-\tau_M z}$ if $\tau_M < +\infty$.

The pair $(\bar{\pi}_0, \bar{\pi}(x))$ satisfies the system

$$\begin{aligned} \bar{\pi}_0 &= \bar{\pi}_0 \pi_{00} + \int_0^{\bar{z}} \pi(x, 0) \bar{\pi}(x) dx \\ \bar{\pi}(x) &= \bar{\pi}_0 \pi(0, x) + \int_0^{\bar{z}} \pi(z, x) \bar{\pi}(z) dz \quad x \in (0, \bar{z}) \\ 1 &= \bar{\pi}_0 + \int_0^{\bar{z}} \bar{\pi}(x) dx. \end{aligned} \quad (21)$$

Proof The proof follows the same path as in Proposition 2.1, but much more simply since the state space is one-dimensional. From any state p one reaches 0 if $\tau < 1$; vice versa from 0 one can reach any state $p \in (0, \bar{z})$ for an appropriate $\tau \in (1, \tau_M)$. The aperiodicity follows from the fact that $\pi_{00} = \mathbb{P}(p^{(k+1)} = 0 | p^{(k)} = 0) > 0$.

It follows that the Markov chain is recurrent, aperiodic, and convergent. Equations (21) come directly from the stationarity condition. \square

Presumably, the result holds also if $\tau_m \geq 1$, but the proof would be more delicate. If $\tau_M \leq 1$, then $\bar{\pi}_0 = 1$, since $P(R_0 \leq 1) = 1$.

Similarly to Subsection 3.3, one may wish to predict $z^{(k)}$ knowing $R_e^{(k)}$; in this case we assume that $p^{(k)}$ is not known, but its distribution is the stationary one $\bar{\pi}$.

One starts by computing the stationary bivariate distribution of $(z^{(k)}, R_e^{(k)})$. If $p^{(k)} = 0$ (with probability $\bar{\pi}_0$), $R_e^{(k)} = \tau_k$ and $z^{(k)} = z(R_e^{(k)})$ where the function $z(\cdot)$ is defined in Proposition 3.1. Hence, the pair $(z^{(k)}, R_e^{(k)})$ has a component on the curve

$(z(R_e), R_e)$ with (one-dimensional) density of R_e given by $\bar{\pi}_0 q_\tau(\cdot)$, with q_τ the marginal distribution of τ .

If $p^{(k)} > 0$, its (defective) density is given by $\bar{\pi}(\cdot)$ and one obtains the bivariate distribution simply by multiplying by $\bar{\pi}(x)$ the formulae presented in Proposition 3.1. Precisely, the (defective) bivariate density over $\{z \in (0, 1), R_e > 1\}$ is

$$\bar{f}(z, R_e) = \begin{cases} 0 & \text{if } -\frac{\ln(1-z)}{z} \geq R_e \\ \int_{1-p(z, R_e)}^1 \bar{\pi}(x) f(z, R_e|x) dx & \text{otherwise} \end{cases} \quad (22)$$

where $f(z, R_e|x)$ is given by (6), and $p(z, R_e)$ is the only solution in $(0, 1)$ of

$$R_e = -\frac{(1-p)}{z} \ln \left(1 - \frac{z}{1-p} \right).$$

From the previous formulae, one obtains, through straightforward computations, the conditional stationary distribution of $z^{(k)}$ given $R_e^{(k)}$. One may note that this distribution has an atom at $z(R_e)$ of probability

$$\frac{\bar{\pi}_0 q_\tau(R_e)}{\bar{\pi}_0 q_\tau(R_e) + \int_0^{z(R_e)} \bar{f}(z, R_e) dz}$$

where $\pi(x, z)$ is given in (16), and a continuous (defective) density over $(0, z(R_e))$

$$\bar{f}(z|R_e) = \frac{\bar{f}(z, R_e)}{\bar{\pi}_0 q_\tau(R_e) + \int_0^{z(R_e)} \bar{f}(z, R_e) dz}$$

4 Numerical illustrations

We start by illustrating the analytical results obtained in the case $r = 2$. We chose 4 different densities q for the pairs (δ_k, τ_k) ; in all cases δ follows the Beta(a, b) distribution, and τ the log-normal(μ, σ^2). In cases 1 and 2 the two variables are independent; in cases 3 and 4 they are negatively correlated. Negative correlation between δ and τ appears realistic, since viral strains may be fitter either because they are more transmissible or because they are able to evade existing immune response.

Parameters and features of the four cases are listed in Table 1, while Figure 1 shows 200 random values from each of the four distributions.

In Figure 2, we show some heat maps of the bivariate distributions of the pair $(z^{(k)}, R_e^{(k)})$ given $p^{(k)} = z^{(k-1)} = p$, computed using Proposition 3.1. To give a better feeling of the distributions we added (small crosses) 500 random draws. It can be seen that either ($p = 0.1$), the support of the distribution is very narrow, or, when the support is wider ($p = 0.5$), the density is concentrated near the curve $R_e = -\log(1-z)/z$.

Parameters	Case 1	Case 2	case 3	Case 4
a	3	4	0.5	3
b	7	6	1.5	7
μ	0.683	1.08	$0.6 - 0.4 \delta$	$0.7 - 0.5 \delta$
σ^2	0.02	0.02	0.02	0.02
$E(\delta)$	0.3	0.4	0.25	0.3
$\sqrt{V(\delta)}$	0.14	0.15	0.25	0.14
$E(\tau)$	2	2.97	1.74	1.83
$\sqrt{V(\tau)}$	0.28	0.42	0.27	0.28
correl. coeff.	0	0	-0.38	-0.32

Table 1 Parameters and features of the 4 distributions used in the simulations. In all cases $\delta \sim \text{Beta}(a, b)$, while $\tau \sim \text{Lognorm}(\mu, \sigma^2)$.

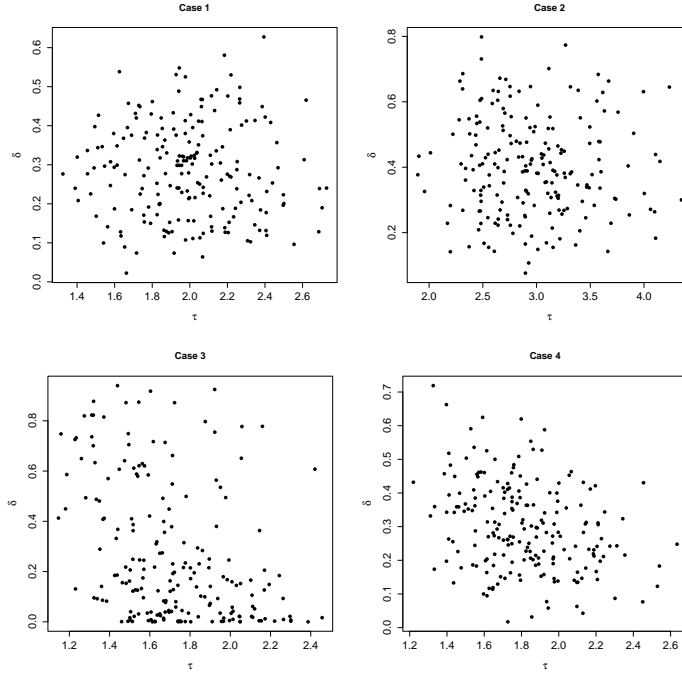


Fig. 1 200 draws of the random variables (τ, δ) for each case described in Table 1

In Figure 3 we show some examples of the distribution of $z^{(k)} = p^{(k+1)}$ given $p^{(k)} = p$ for two different values of p ; clearly, these correspond to the transition probabilities, obtained from Proposition 3.2. In the same pictures we also show the distribution assuming that also $R_e^{(k)}$ is known and equal to 1.6. It can be seen that the distributions of $z^{(k)}$ is quite different depending on the value of p . Knowledge of $R_e^{(k)}$ hence greatly reduces uncertainty (i.e. randomness) of $z^{(k)}$. As a consequence, a prediction of the final attack ratio after observing the initial growth in an epidemic will greatly improve

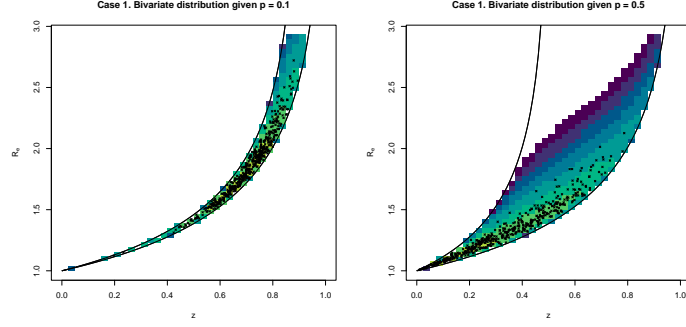


Fig. 2 Examples of the bivariate distributions of $(z^{(k)}, R_e^{(k)})$ given $p^k = z^{(k-1)} = p$ computed using Proposition 3.1 for $p = 0.1$ (left panel) or $p = 0.5$ (right panel) in case 1 of Table 1.

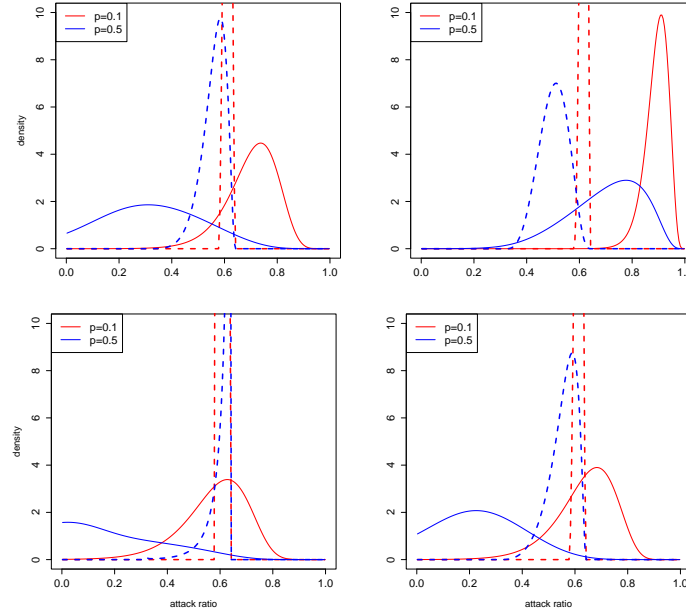


Fig. 3 The distributions of $z^{(k)}$ given $p^{(k)} = z^{(k-1)}$ (solid curves), and given the pair $(R_e^{(k)}, p^{(k)})$ (dashed curves). The red curves are computed with $p^{(k)} = p = 0.1$, the blue ones with $p = 0.5$. The conditional distributions are based on $R_e^{(k)} = 1.6$. The four panels correspond to the four distributions of the pairs (δ_k, τ_k) as shown in Table 1.

the precision of a prediction of the final attack ratio $z^{(k)}$ as compared to only knowing the attack ratio of the year before.

In Figure 4 we show (black curves) the stationary distribution of $z^{(k)}$ computed through simulation of 20,000 values, smoothed through the `density` function of R. These are compared to the conditional distributions of $z^{(k)}$ given $R_e^{(k)}$ for $R_e^{(k)} = 1.6$ and 1.3, being common values for seasonal influenza estimates (e.g. [8]).

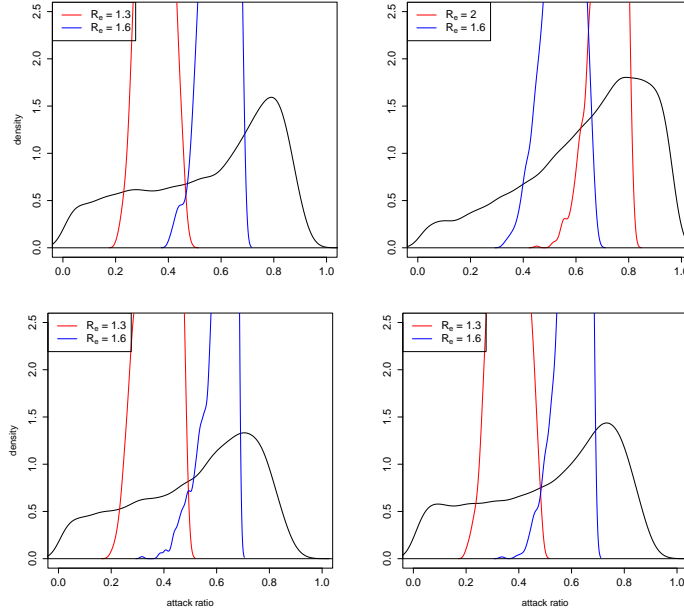


Fig. 4 Estimates of the stationary distribution (black curve) of the attack ratio $z^{(k)}$, and of the conditional distributions of $z^{(k)}$ given $R_e^{(k)} = R_e$ (red and blue curves); values used for R_e in the legend. The four panels correspond to the four distributions of the pairs (δ_k, τ_k) as shown in Table 1.

Note that the density shown for the stationary distributions (as well the transition probabilities shown in Figure 3) are defective, in the sense that the stationary distribution also has probability mass at $z = 0$. The value of π_0 is estimated at 0.25, 0.06, 0.32 and 0.28 in the four cases illustrated in the figure.

Note also that we showed in Subsection 3.4 that the conditional distributions of $z^{(k)}$ given $R_e^{(k)} = R_e$ has an atom at $z(R_e)$; this is not so clear in Figure 4, because it is an empirical estimate of the distribution from a simulation through the function **density**, and because we selected the values of k $R_e^{(k)}$ was close to R_e and not exactly equal to R_e (which is practically impossible). For these reasons, the atom of $z(R_e)$ looks blurred in the Figure.

In Figure 5 we show the distribution of $(R_e^{(k)}, z^{(k)})$ along the simulations. It can be seen that in all cases, the points are rather close to the curve $R_e = -\log(1 - z)/z$; in other words, if the model were correct, predicting the attack ratio from the final size equation using the value of $R_e^{(k)}$ (without using the knowledge of $z^{(k-1)}$) would give a rather good result.

We have also considered the more realistic case of $r = 10$, i.e. immunity is completely lost 10 years after the last infection; in this case the state space has dimension 17, and thus the transition probabilities are very complicated to compute analytically and visualize.

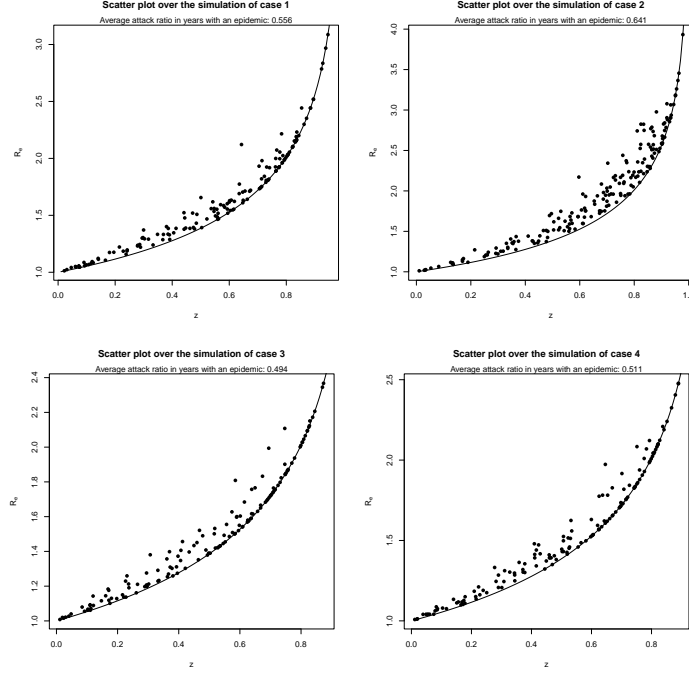


Fig. 5 Bivariate graphs of $(R_e^{(k)}, z^{(k)})$ found along the simulations of the model with $r = 2$. The curves are the functions $R_e = -\log(1-z)/z$ above which the bivariate distribution always lies, as shown in Section 3.

Therefore, when $r = 10$, we limited ourselves to simulations of the model. In Figure 6 we repeat the same procedure used to obtain Figure 4. In this case the stationary distribution is multi-dimensional, and we show estimates of the marginal distribution of the final size $z^{(k)}$. If one compares Figure 6 to Figure 4, one sees that the densities are shifted to the left with $r = 10$, i.e. when immunity lasts longer the average attack ratio is lower. On the other hand, once $R_e^{(k)}$ is known, the estimates of $z^{(k)}$ are similar for $r = 2$ and $r = 10$.

In Figure 7 we show, analogously to Figure 5, the pairs $(R_e^{(k)}, z^{(k)})$.

Comparing the results obtained in the case $r = 10$ with the corresponding ones for $r = 2$, one notices that the values of the pairs $(R_e^{(k)}, z^{(k)})$ are even closer to the curve $R_e = -\log(1-z)/z$ than for $r = 2$. The fraction of years having an epidemic outbreak is similar for $r = 10$ and $r = 2$, except for case 3, while the average size of the epidemics is lower when $r = 10$ than when $r = 2$, again less so in case 2, in which immunity drift is generally larger.

5 Discussion

In the paper we have introduced a stochastic epidemic model for seasonal outbreaks (of e.g. influenza), where the virus has a random genetic drift and a new random transmissibility each year. Given the community immunities coming into season k ,

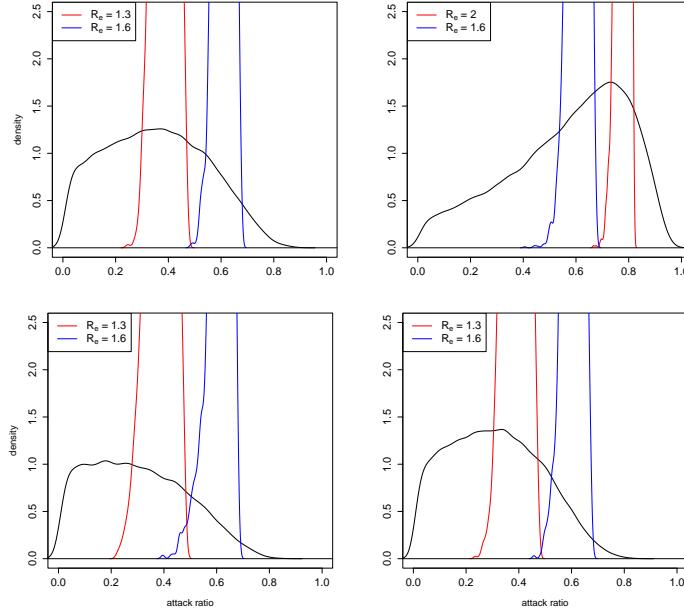


Fig. 6 Estimates of the marginal distribution (black curve) of the attack ratio $z^{(k)}$, and of the conditional distributions of $z^{(k)}$ given $R_e^{(k)} = R_e$ (red and blue curves), in the stationary distribution of the model with $r = 10$; values used for R_e in the legend. The four panels correspond to the four distributions of the pairs (δ_k, τ_k) as shown in Table 1.

the epidemic outcome, described by the effective reproduction number $R_e^{(k)}$ and final attack ratio $z^{(k)}$, is determined by the drift and transmissibility of this year's virus strain. The novelty of our model lies in considering the yearly drift as random, varying between years, and to also allow for transmissibility to vary between years.

The process of community immunities coming into next year's season, is shown to be a Markov chain having a stationary distribution. The case where immunity only lasts one year is studied in more detail. For this situation, we obtain analytical results for the transition distribution of the Markov chain, its stationary distribution, and also the conditional distribution of the final attack ratio, given the early epidemic growth (corresponding to $R_e^{(k)}$) and incoming immunity.

The hope is that our model can contribute to the understanding of how and why influenza outbreaks vary in size and initial growth between years, and also potentially help to improve precision in predicting the severity of an influenza season based on observing the start of the outbreak and knowing earlier years outcome.

Our model of course has several simplifications compared to empirical seasonal epidemics. The perhaps most important extensions for making the model more realistic, are a) to consider more than one circulating strain: influenza currently has three main circulating strains (AH1N1, AH3N2 and B) with separate attack ratios varying between years, and with interactions between the circulating strains (e.g. [9]),

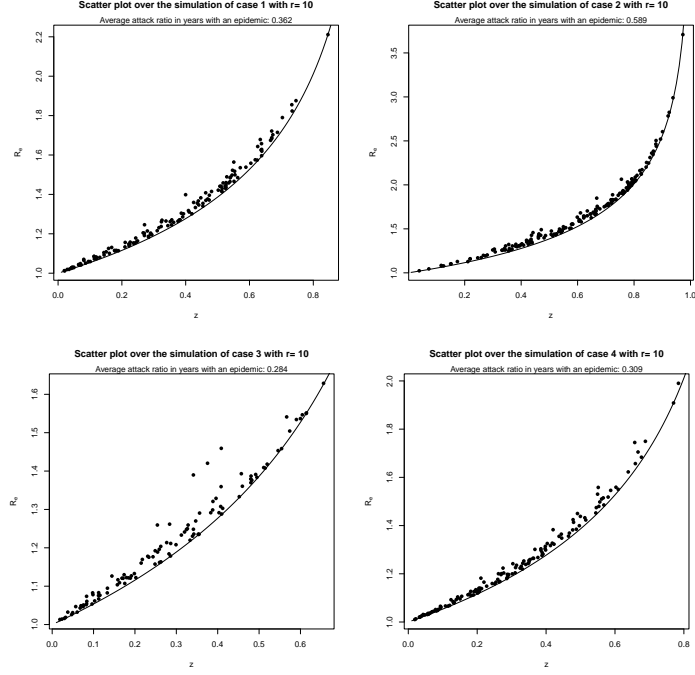


Fig. 7 Bivariate graphs of $(R_e^{(k)}, z^{(k)})$ found along the simulations of the model with $r = 10$. The curves are the functions $R_e = -\log(1 - z)/z$ above which the bivariate distribution always lies, as shown in Section 3.

and b) to allow for heterogeneities in the community, for example in terms of partial immunities and how these affect infectivity.

In our illustrations we assume that the drift and transmissibility of different years are independent. Our framework however allows for dependencies between years, for example that large genetic drifts rarely happen repeatedly consecutive years, and this can be incorporated, still keeping the Markov property in our model.

From an applied point of view, it would be interesting to try to fit the model to seasonal outbreaks over a range of years, and to analyse how the fitted stationary distribution, transition distribution, and early season predictions fit to data. It would also be of interest to analyse what underlying distribution of the drift and transmissibility that fits best to observed seasonal outbreak data, which may then be compared to other estimates of yearly drift and transmissibility changes.

Of more mathematical curiosity, it would be interesting to see if a model allowing longer lasting immunity (larger r) is more flexible than when immunity only lasts one year. It would also be interesting to obtain qualitative results regarding concentration of randomness of the final size coming from knowing prior immunity and/or the initial growth rate, in line with what was illustrated numerically in Section 4.

Acknowledgements

T.B. is grateful to the Swedish Research Council (grant 2020-0474) for financial support. This work was initiated when T.B. visited University of Trento for 4 weeks in the spring 2025. He is grateful to the hospitality of the Department of Mathematics in Trento. A.P. thanks the Italian Ministry for University and Research (MUR) through the PRIN 2020 project “Integrated Mathematical Approaches to Socio-Epidemiological Dynamics” (No. 2020JLWP23). A.P. is member of the “Gruppo Nazionale per l’Analisi Matematica e le sue Applicazioni” (GNAMPA) of the “Istituto Nazionale di Alta Matematica” (INdAM).

Data availability statements

The paper does not analyse any data. Simulation programmes can be requested from the authors.

References

- [1] O Diekmann, H Heesterbeek, T Britton. (2013). *Mathematical tools for understanding infectious disease dynamics*. Princeton UP.
- [2] Andreasen V. (2003). Dynamics of annual influenza A epidemics with immunoselection. *J Math Biol.* **46**: 504-536. <https://doi.org/10.1007/s00285-002-0186-2>.
- [3] G Chowell, M A Miller and C Viboud. (2007). Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. *Epidemiol. Infect.* **136**, 852-864. doi:10.1017/S0950268807009144
- [4] Durrett R. (2019). *Probability: theory and examples*. Cambridge UP.
- [5] Jacod J, Protter P (2000) *Probability essentials*. Springer.
- [6] Kucharski AJ, Gog JR and Andreasen V (2016). Capturing the dynamics of pathogens with many strains. *J Math Biol.* **72**: 1-24. <https://doi.org/10.1007/s00285-015-0873-4>.
- [7] Roberts MG, Hickson RI and McCaw JM (2024). How immune dynamics shape multi-season epidemics: a continuous-discrete model in one dimensional antigenic space. *J Math Biol.* **88**: 48. <https://doi.org/10.1007/s00285-024-02076-x>.
- [8] F Trentini, E Pariani, A Bella, G Diurno, L Crottogini, C Rizzo, S Merler and M Ajelli. (2022). Characterizing the transmission patterns of seasonal influenza in Italy: lessons from the last decade. *BMC Public Health.* **22**:19 <https://doi.org/10.1186/s12889-021-12426-9>
- [9] W Yang, E H Y Lau and B J Cowling. (2020). Dynamic interactions of influenza viruses in Hong Kong during 1998-2018. *PLoS Comp Biol.* **16**(6): e1007989. <https://doi.org/10.1371/journal.pcbi.1007989>.