Spatial Spread of Epidemic Diseases in Geographical Settings: Seasonal Influenza Epidemics in Puerto Rico

Pierre Magal [1,2], G.F. Webb [3], Yixiang Wu [3]

1 Univ. Bordeaux, IMB, UMR 5251, Talence F-33400, France
2 CNRS, IMB, UMR 5251, Talence F-33400, France
3 Mathematics Department, Vanderbilt University, Nashville, TN

Abstract

A deterministic model is developed for the spatial spread of an epidemic disease in a geographical setting. The model is focused on outbreaks that arise from a small number of infected individuals in sub-regions of the geographical setting. The goal is to understand how spatial heterogeneity influences the transmission dynamics of susceptible and infected populations. The model consists of a system of partial differential equations with a diffusion term describing the spatial spread of an underlying microbial infectious agent. The model is applied to simulate the spatial spread of the 2016-2017 seasonal influenza epidemic in Puerto Rico. In this simulation, the reported case data from the Puerto Rican Department of Health are used to implement a numerical finite element scheme for the model. The model simulation explains the geographical evolution of this epidemic in Puerto Rico, consistent with the reported case data.

1 Introduction

Epidemic outbreaks evolve in geographical regions with considerable variability in spatial locations. This spatial variability is important in understanding the impact of public health policies and interventions in controlling these epidemics. A major difficulty in developing models to describe spatial variability in epidemics is accounting for the movement of people in spatial contexts. Many approaches to produce such descriptions have been developed, including individual based models, network models, stochastic models, and partial differential equations models. Individual based, network, and stochastic models employ societal data of human movement and interaction to simulate human behavior at spatial and temporal levels based on probabilistic assumptions ([6], [14], [18], [22], [23], [27], [28], [29], [31], [38], [34], [40], [46], [47], [50], [54], [61], and many others). These models sometimes require intensive informational input, as well as intensive computational output. Partial differential equations models offer an alternative approach, with advantages for both informational input and computational output.

Partial differential equations models have been developed by many authors to study the spread of diseases in spatial settings, with diffusion terms used to describe the movement of both susceptible and infectious individuals ([1], [13], [19], [24], [37], [57], [59], and many others). Some studies have used diffusion to describe only infectious populations ([12], [20], [25], [26], [33], [53], [35]). Our objective here is two-fold: (1) to analyze a reaction-diffusion epidemic model, with diffusion modeling only the infectious population (viewed indirectly as the movement of an infectious agent through the susceptible population); and (2) to apply the model to the 2016-2017 seasonal influenza epidemic in Puerto Rico, with comparison to reported case data.

Seasonal influenza epidemics recur annually during the cold half of the year in each hemisphere. Each annual flu season is normally associated with a major influenza-virus subtype. The associated subtype typically changes each year, due to development of immunological resistance to a previous year's strain through exposure and vaccinations, and mutational changes in previously dormant viral strains. The exact mechanism behind the seasonal nature of influenza outbreaks is unknown (https://en.wikipedia.org/wiki/Flu-season). We refer to [11, 15, 17, 32] for studies on the spatial-temporal transmission of influenza and to [39, 48, 62] for studies on the seasonality of influenza.

Our objective for the simulation of the 2016-2017 Puerto Rico influenza epidemic is to match our model to reported case data, focusing on epidemic duration, attack rates, turning points, final size, and geographical features. Our simulations indicate that epidemic duration depends strongly on the local depletion of the susceptible population to levels that no longer sustain transmission of the infection. This depletion happens rapidly in local geographical regions, while the general level of the epidemic lasts longer in larger geographical regions. These geographical distinctions offer opportunity for public health policy interventions to control epidemic severity.

This paper is organized as follows: In Section 2 we formulate a general deterministic model for the evolution of an epidemic outbreak in a spatial domain. In Section 3 we state and prove theorems for our deterministic model of a spatial epidemic. In Section 4 we specify the model to the 2016-2017 seasonal influenza epidemic in Puerto Rico. In Section 5 we discuss implications of our results for public health policies for controlling an epidemic outbreak.

2 A General Deterministic Spatial Epidemic Model

In most applications of reaction-diffusion models to epidemics, diffusion does not provide a realistic description of the way people move in societal settings. Diffusion provides only an averaging process that cannot account for the extreme spatial and temporal heterogeneity of human movement. Diffusion alters the home-base of individuals, which is not realistic during the time-evolution of an epidemic. We argue, alternatively, that the spatial movement of the micro-organisms causing the epidemic, rather than the spatial movement of humans, is a realistic way to account for epidemic spatial development. The movement of the infectious agent can be viewed indirectly, as the movement of infectious individuals, described with diffusion processes.

It is clear that the contributions of local-distance and long-distance transmission are both involved in the spatial evolution of epidemics Arino [5], Arino and Khan [4], Charu [16]. In [28], however, it is argued that for the 2009 H1N1 influenza epidemic, local transmissions were of greater importance than distant transmissions, as outbreaks in proximate communities resulted in successful infection chains, whereas, distant transmissions died out after a small number of generations. The underlying assumption is that most infections occur close to home-base of infectious individuals, which spread to nearby susceptible individuals. In our model, the diffusion of infected individuals occurs only during their time-span of infectiousness, which is typically only a few days. We claim that this movement, while not significant for them, is significant for the virus they spread to susceptible individuals nearby. Our simplifying assumption is that susceptible individuals do not move, but have highest infection probability at their home-base, to which they regularly return.

Our model has the following formulation: $\Omega \subset \mathbb{R}^2$ is a bounded domain, $S(t, \mathbf{x})$ and $I(t, \mathbf{x})$ are the spatial densities at location $\mathbf{x} \in \Omega$ and at time t of susceptible and infected individuals,

respectively.

$$
\begin{aligned}
\frac{\partial}{\partial t}S(t,\mathbf{x}) &= -\frac{\tau(\mathbf{x})\,I(t,\mathbf{x})^p}{1+\kappa(\mathbf{x})\,I(t,\mathbf{x})^q}S(t,\mathbf{x}),\ \mathbf{x}\in\Omega,\ t>0\\
\frac{\partial}{\partial t}I(t,\mathbf{x}) &= \alpha\Delta I(t,\mathbf{x})+\frac{\tau(\mathbf{x})\,I(t,\mathbf{x})^p}{1+\kappa(\mathbf{x})\,I(t,\mathbf{x})^q}S(t,\mathbf{x})-\lambda(\mathbf{x})I(t,\mathbf{x}),\ \mathbf{x}\in\Omega,\ t>0\\
\frac{\partial}{\partial \eta}I(t,\mathbf{x}) &= 0, \mathbf{x}\in\partial\Omega,\ t>0\\
S(0,\mathbf{x}) &= S_0(\mathbf{x}),\ I(0,\mathbf{x})=I_0(\mathbf{x}),\ \mathbf{x}\in\Omega
\end{aligned}
$$

where α is the diffusion parameter for infected individuals, $\tau(\mathbf{x}),\kappa(\mathbf{x}),p$ and q are transmission parameters, and $\lambda(\mathbf{x})$ is the removal rate of infected individuals. The transmission rate has nonlinear incidence form (Hethcote [30], Liu [38], Ruan [52]), where $\tau I(t,\mathbf{x})^p$ measures the force of infectiousness and $1/(1+\kappa I(t,\mathbf{x})^q)$ measures reduced infectiousness resulting from behavioral change as the number of infected individuals increases ($1\le p\le q+1$). Also, $\kappa(\mathbf{x}),\tau(\mathbf{x})$, $\lambda(\mathbf{x})$ are positive continuous functions on $\overline{\Omega}$, and the initial data S_0 and I_0 are nonnegative continuous functions on $\overline{\Omega}$. The theoretical analysis of the model is given in the next section.

3 Analysis of the model

In this section we investigate the long term behavior of the system. One may observe that some special case of our model have been considered by Britton [7], Murray [49] and more recently by Ducrot and Gilletti [8]. Since they consider some special case of our system they can reduce the system to a single equation. In the general case we were not able to extend this kind of idea. Therefore we need to provide an alternative approach.

We start with the ordinary differential equation case which corresponds for (2.1) to the case where the distributions $x\to I(t,x)$ and $x\to S(t,x)$ (as well as the parameters) are constants.

Theorem 3.1 *Let $\tau,\kappa,\lambda,p,q>0$ with $1\le p\le q+1$, and let $S_0,I_0>0$. There exists a unique solution $S(t)\ge 0$, $I(t)\ge 0$, satisfying $S(0)=S_0$, $I(0)=I_0$, and*

$$
S'(t) = -\frac{\tau\,I(t)^p}{1+\kappa\,I(t)^q}S(t), \tag{3.1}
$$

$$
I'(t) = \frac{\tau\,I(t)^p}{1+\kappa\,I(t)^q}S(t)-\lambda I(t). \tag{3.2}
$$

Let $R_0=\tau I_0^{p-1}S_0/(\lambda(1+\kappa I_0^q))$. If $R_0<1$, then $S(t)$ decreases to a limiting value $S_\infty>0$ and $I(t)$ decreases to 0. If $R_0>1$, then $S(t)$ decreases to a limiting value $S_\infty>0$ and $I(t)$ first increases, then decreases to 0.

Proof. Add (3.1) and (3.2) and integrate over $(0,t)$ to obtain

$$
0\le S(t)+I(t)+\lambda\int_0^t I(s)ds = S_0+I_0. \tag{3.3}
$$

The existence of a unique nonnegative solution on $[0,\infty)$ follows from standard theory. Since $S'(t)\le 0$, $S(t)$ converges to a limt $S_\infty\ge 0$. Also, $S(t)$ and $I(t)$ are bounded on $[0,\infty)$, $I'(t)$ is bounded on $[0,\infty)$, and

$$\int_0^\infty I(t)dt < \infty,$$

which implies that $\lim_{t\to\infty} I(t) = 0$. Noticing $1 \le p \le q+1$, a simple calculation shows that

$$\frac{z^p}{1+\kappa\, z^q} \le \max(1, 1/\kappa) z,\ z \ge 0. \tag{3.4}$$

Thus,

$$\int_0^\infty \frac{\tau\, I(t)^p}{1+\kappa\, I(t)^q} dt \le \tau \max(1, 1/\kappa) \int_0^\infty I(t)dt \ < \ \infty. \tag{3.5}$$

Divide both sides of (3.1) by $S(t)$ and integrate over $(0, t)$ to obtain

$$\ln\left(\frac{S(t)}{S_0}\right) = -\int_0^t \frac{\tau\, I(s)^p}{1+\kappa\, I(t)^q} ds \tag{3.6}$$

which implies

$$S_\infty = S_0 \exp\left(-\int_0^\infty \frac{\tau\, I(t)^p}{1+\kappa\, I(t)^q} dt\right) \ne 0.$$

Then, to show that $I(t)$ can have at most one peak, observe from (3.2)

$$I''(t) = \Bigg(\Big(1+\kappa I(t)^q\Big)\Big(\tau p I(t)^{p-1} I'(t) S(t) + \tau I(t)^p S'(t)\Big)$$

$$- \Big(\tau I(t)^p S(t)\Big)\Big(\kappa q I(t)^{q-1} I'(t)\Big)\Bigg) \Big/ \Big(1+\kappa I(t)^q\Big)^2 - \lambda I'(t).$$

If $I'(\bar{t}) = 0$, then

$$I''(\bar{t}) = \frac{\tau I(\bar{t})^p S'(\bar{t})}{1+\kappa I(\bar{t})^q} < 0,$$

which implies $I(t)$ is concave down wherever $I'(\bar{t}) = 0$.

Rewrite (3.2) as

$$I'(t) = \lambda\left(\frac{\tau\, I(t)^{p-1} S(t)}{1+\kappa\, I(t)^q} - 1\right) I(t).$$

Then we can see that $I(t)$ decreases at $t = 0$ if $R_0 < 1$ and increases at $t = 0$ if $R_0 > 1$. So the claim on $I(t)$ follows from the fact that $I(t)$ converges to zero and has at most one peak. ■

Remark 3.2 *If $p = 1$ and $\kappa = 0$, one can combine (3.3) and (3.6) to obtain*

$$S_\infty + \frac{\lambda}{\tau} \ln\left(\frac{S_\infty}{S_0}\right) = S_0 + I_0.$$

Theorem 3.3 *Let Ω be a bounded domain in $\mathbb{R}^n$ with smooth boundary $\partial\Omega$. Let $\alpha, \tau, \kappa, p, q$ be positive constants with $1 \le p \le q$ and $\kappa > 0$, let $\lambda \in C_+(\overline{\Omega})$ with $\lambda(\mathbf{x}) \ge \lambda_0 > 0$ for all $\mathbf{x} \in \Omega$, and let S_0, $I_0 \in L^1_+(\Omega)$ with $S_0 \ne 0$ and $I_0 \ne 0$ and there exists a constant $S_0^+ > 0$ such that*

$$S_0(x) \le S_0^+,\ \textit{for a.e. } \mathbf{x} \in \Omega.$$

Then, there exists unique $S(t,\cdot), I(t,\cdot) : [0,\infty) \to L^1_+(\overline{\Omega})$ *satisfying*

$$\frac{\partial}{\partial t}S(t,\mathbf{x}) = -\frac{\tau\, I(t,\mathbf{x})^p}{1+\kappa\, I(t,\mathbf{x})^q}S(t,\mathbf{x}),\ \mathbf{x}\in\Omega,\ t>0 \tag{3.7}$$

$$\frac{\partial}{\partial t}I(t,\mathbf{x}) = \alpha\Delta I(t,\mathbf{x}) + \frac{\tau\, I(t,\mathbf{x})^p}{1+\kappa\, I(t,\mathbf{x})^q}S(t,\mathbf{x}) - \lambda(\mathbf{x})I(t,\mathbf{x}),\ \mathbf{x}\in\Omega,\ t>0 \tag{3.8}$$

$$\frac{\partial}{\partial \eta}I(t,\mathbf{x}) = 0,\ \mathbf{x}\in\partial\Omega,\ t>0 \tag{3.9}$$

$$S(0,\mathbf{x}) = S_0(\mathbf{x}),\ I(0,\mathbf{x}) = I_0(\mathbf{x}),\ \mathbf{x}\in\Omega. \tag{3.10}$$

Moreover

$$\lim_{t\to\infty} S(t,\cdot) = S_\infty(\cdot) \geq 0,\ \lim_{t\to\infty} I(t,\cdot) = 0 \text{ in } L^1(\Omega), \tag{3.11}$$

with

$$\iint_\Omega S_\infty(\mathbf{x})dx > 0,$$

whenever there exists a constant $\delta > 0$ *such that* $S_0(\mathbf{x}) \geq \delta$ *for a.e.* $\mathbf{x}\in\Omega$.

Proof. The map $f(z) := \frac{z^p}{1+\kappa\, z^q}$ is Lipschitz continuous on $[0,\infty)$ whenever $1 \leq p \leq q+1$ and $\kappa > 0$, because its derivative

$$f'(z) = \frac{pz^{p-1} + \kappa(p-q)z^{p+q-1}}{(1+\kappa z^q)^2}$$

is bounded on $[0,\infty)$. Moreover since $p \leq q$ and $\kappa > 0$, the map f is bounded. It follows that the map $(S,I) \to f(I)S$ is Lipschitz continuous from M into $L^1(\Omega)$, where

$$M := \left\{(S,I) \in L^1_+(\Omega)^2 : S(\mathbf{x}) \leq S_0^+,\ \text{for almost every } \mathbf{x}\in\Omega\right\}.$$

Indeed we have

$$\|f(I)S - f(\widehat{I})\widehat{S}\|_{L^1} \leq \|f(I)S - f(\widehat{I})S\|_{L^1} + \|f(\widehat{I})S - f(\widehat{I})\widehat{S}\|_{L^1}$$

hence

$$\|f(I)S - f(\widehat{I})\widehat{S}\|_{L^1} \leq \left[\|f\|_{\mathrm{Lip}}S^+ + \|f\|_\infty\right]\left[\|I-\widehat{I}\|_{L^1} + \|S-\widehat{S}\|_{L^1}\right].$$

Therefore by using standard argument, the system (3.7)-(3.10) generates a unique maximal continuous semiflow in M.
By using the standard blowup conditions we can see that for each $(S_0, I_0) \in M$ the blowup time is $+\infty$ because $t \to S(t,x)$ is positive and decreasing and by adding (3.7) and (3.8) and integrate over $(0,t)$ and Ω to obtain

$$\iint_\Omega (S(t,\mathbf{x}) + I(t,\mathbf{x}))d\mathbf{x} + \iint_\Omega \left(\int_0^t \lambda(\mathbf{x})I(s,\mathbf{x})ds\right)d\mathbf{x} = \iint_\Omega (S_0(\mathbf{x}) + I_0(\mathbf{x}))d\mathbf{x}, \tag{3.12}$$

which implies

$$\iint_\Omega I(t,\mathbf{x})d\mathbf{x} \leq \iint_\Omega (S_0(\mathbf{x}) + I_0(\mathbf{x}))d\mathbf{x}.$$

Therefore the solution starting in M remains in M and is bounded. Moreover by using again (3.12), we deduce that

$$\iint_\Omega \left(\int_0^\infty \lambda(\mathbf{x}) I(t,\mathbf{x}) dt \right) d\mathbf{x} \leq \iint_\Omega (S_0(\mathbf{x}) + I_0(\mathbf{x})) d\mathbf{x}.$$

Therefore, by using (3.4) and the fact that $\lambda(\mathbf{x}) \geq \lambda_0 > 0$, we deduce that

$$\iint_\Omega \left(\int_0^\infty \frac{\tau\, I(t,\mathbf{x})^p}{1+\kappa\, I(t,\mathbf{x})^q} dt \right) d\mathbf{x} \leq \tau\, \max(1, 1/\kappa) \iint_\Omega \left(\int_0^\infty I(t,\mathbf{x}) \mathbf{dt} \right) \mathbf{dx} < \infty.$$

As in the ODE case, (3.7) implies that for a.e. $\mathbf{x} \in \Omega$, $\frac{\partial}{\partial t} S(t,\mathbf{x}) \leq 0$ and since the norm of L^1 is additive on L^1_+ (i.e. by using the fact that $\|S(t,.)\|_{L^1} = \|S(t,.) - S(t+l,.)\|_{L^1} + \|S(t+l,.)\|_{L^1}, \forall t, l \geq 0,$). We deduce that for each increasing sequence $t_n \in (0,\infty) \to \infty$ as $n \to \infty$, the sequence $n \to S(t_n,.)$ is a Cauchy in $L^1(\Omega)$. Therefore

$$\lim_{t\to\infty} S(t,\mathbf{x}) = S_\infty(\mathbf{x}) \geq 0 \text{ in } L^1(\Omega).$$

Integrate (3.7) over t to obtain for a.e. $\mathbf{x} \in \Omega$,

$$\ln\left(S_\infty(\mathbf{x})\right) = \ln\left(S_0(\mathbf{x})\right) - \int_0^\infty \frac{\tau\, I(s,\mathbf{x})^p}{1+\kappa\, I(s,\mathbf{x})^q} ds.$$

Assume that $S_0(\mathbf{x}) \geq \delta$ for a.e. $\mathbf{x} \in \Omega$ for some $\delta > 0$. Then we have

$$|\Omega|^{-1} \iint_\Omega \ln(S_\infty(\mathbf{x})) dx = |\Omega|^{-1} \iint_\Omega \ln(S_0(\mathbf{x})) dx - |\Omega|^{-1} \iint_\Omega \int_0^\infty \frac{\tau\, I(s,\mathbf{x})^p}{1+\kappa\, I(s,\mathbf{x})^q} ds dx$$

and since the function $\ln(\mathbf{x})$ is concave, by using the Jensen's inequality we obtain

$$\ln(|\Omega|^{-1} \iint_\Omega S_\infty(\mathbf{x}) dx) \geq |\Omega|^{-1} \iint_\Omega \ln(S_\infty(\mathbf{x})) dx$$

it follows that

$$\iint_\Omega S_\infty(\mathbf{x}) dx > 0.$$

For $S_0, I_0 \in L^1_+(\Omega)$, define the ω-limiting set of (S_0, I_0) in $[L^1(\Omega)]^2$ as

$$\{(u,v) \in [L^1(\Omega)]^2 : (S(t_n,\cdot), I(t_n,\cdot)) \to (u,v) \text{ in } [L^1(\Omega)]^2 \text{ for some } \{t_n\}\}.$$

The ω-limiting set of (S_0, I_0) is bounded in $L^1(\Omega)$. Since $S(t,\cdot)$ is convergent in $L^1(\Omega)$, $\{S(t,\cdot) : t \geq 0\}$ has a compact closure in $L^1(\Omega)$. Since the linear operator semigroup generated by the Laplacian with Neumann boundary conditions is compact in $L^1(\Omega)$, the nonlinear term in (3.8) is bounded in t, and since $\lambda_0 > 0$, $\{I(t,\cdot) : t \geq 0\}$ has compact closure in $L^1(\Omega)$ (see Martin [45, Proposition 5.4], Webb [60] or Magal and Thieme [43]). Thus, the ω-limiting set of (S_0, I_0) is non-empty in $L^1(\Omega)$.

To prove $\lim_{t\to\infty} I(t) = 0$, define $V(S,I)(t) = \iint_\Omega (S(t,\mathbf{x}) + I(t,\mathbf{x})) d\mathbf{x}$ and add (5) and (6) to obtain

$$\dot{V}(S,I)(t) = -\iint_\Omega \lambda(\mathbf{x}) I(t,\mathbf{x}) d\mathbf{x} \leq 0.$$

By LaSalle's [36] invariance principle $(S(t), I(t))$ converges to $(S_\infty, 0)$ in $L^1(\Omega)$, since the maximal invariant subset of $\{(S,I) \in M : \dot{V}(S,I) = 0 \text{ and } S = S_\infty\}$ is $(S_\infty, 0)$. ■

Remark 3.4 *If we assume $1 \le p \le q+1$ and $S_0, I_0 \in L^\infty_+(\Omega)$ in Theorem 3.3, then (3.7)-(3.10) has a unique solution $(S(t,\cdot), I(t,\cdot)) : [0,\infty) \to L^\infty_+(\Omega)$ and the convergence $I(t,\cdot) \to 0$ is in $L^\infty(\Omega)$. Indeed, the existence of local solution is standard, and the solution is global if it does not blow up in finite time. Since $S(t,\mathbf{x})$ is decreasing in t, it is uniformly bounded in $L^\infty(\Omega)$ by $\|S_0\|_{L^\infty}$. Equation (3.8) can be written as*

$$\frac{\partial}{\partial t} I(t,\mathbf{x}) = \alpha \Delta I(t,\mathbf{x}) + g(t,\mathbf{x}) I(t,\mathbf{x}), \tag{3.13}$$

where $g(t,\mathbf{x})$ is uniformly bounded in $L^\infty(\Omega)$ since $p \le q+1$. (3.13) can be viewed as a linear equation, and the solution does not blow up. Therefore, the solution of (3.7)-(3.10) exists globally in $L^\infty(\Omega)$.

To see the convergence of $I(t,\mathbf{x})$ in $L^\infty(\Omega)$, by (3.11), it suffices to show that $\{I(t,\cdot)\}_{t>1}$ is precompact in $L^\infty(\Omega)$. Since $\|I(t,\cdot)\|_{L^1}$ is uniformly bounded for $t \ge 0$, by (3.13) and [2, Theorem 3.1], we have that $\|I(t,\cdot)\|_{L^\infty}$ is uniformly bounded for $t \ge 0$. The compactness of the orbit $\{I(t,\cdot)\}_{t>1}$ follows from the fact that the semigroup generated by the Laplacian with Neumann boundary conditions is precompact in $C(\bar{\Omega})$ (see Martin [45, Proposition 5.4], Webb [60] or Magal and Thieme [43]).

4 The 2016-2017 Influenza Epidemic in Puerto Rico

The island of Puerto Rico consists of 76 municipalities, with total population of almost 3,500,000, in a geographical region of approximately 170 km by 60 km. Four major population centers are the municipalities San Juan (northeast - population 395,000), Ponce (south central - population 262, 000), Arecibo (northwest - population 193,000), and Mayagűez (far west - population 89,000) (see Figure 1).

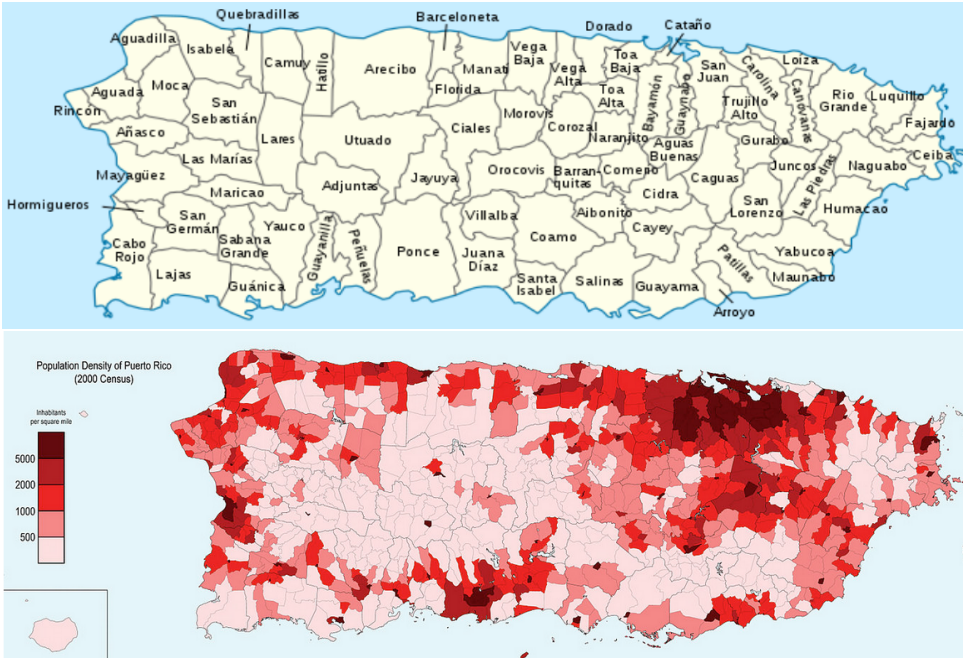

Figure 1: *Top. The 76 municipalities in Puerto Rico (wikipedia.org). Bottom. The population density of Puerto Rico (wikipedia.org).*

Our numerical simulations use the finite element method package in *Matlab*. The first step is to describe the geographical region of Puerto Rico in *Matlab*. The boundary data of Puerto Rico are latitudes and longitudes obtained from *Mathematica* using CountryData["PuertoRico", "SchematicPolygon"], which forms a polygon with 66 points. The latitude bounds for Puerto Rico are $\{17.9, 18.5\}$, the longitude bounds are $\{-67.3, -65.3\}$. The latitudes and longitudes are converted into kilometers for convenience. The boundary data is then used to generate a mesh with 23,772 mesh nodes.

We assume that everyone is susceptible at the beginning of the epidemic season, and therefore $S_0(\mathbf{x})$ equals the population density. This assumption is a simplification, since a significant, but unknown, fraction of the population has immunity due to prior infections, vaccination, or other reasons. We obtain the population density data and the geographical location data (latitude and longitude) for the center of each municipality from *Mathematica* using the function AdministrativeDivisionData. Then, in *Matlab*, we use the interpolation function scatteredInterpolant to calculate the population density at the 23772 mesh nodes (see Figure 2 for a heat map of $S_0(\mathbf{x})$). This method generates a total population of about 4,000,000, which slightly over-estimates the total population of Puerto Rico in 2016-2017. This over-estimation, however, is compensated by the fact that Puerto Rico attracts about 4 million visitors a year (https://en.wikipedia.org/wiki/Tourism.in.Puerto.Rico).

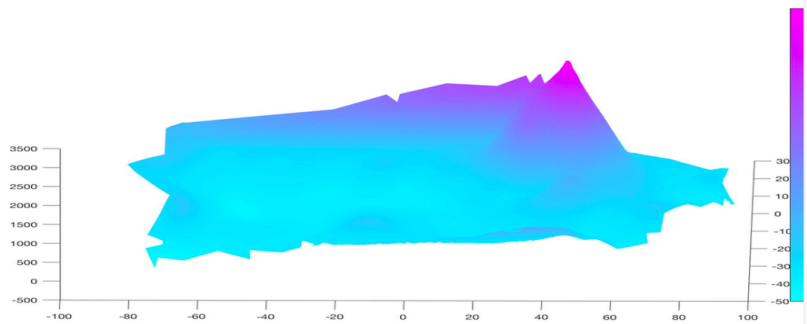

Figure 2: *The population density of the initial susceptible population* $S_0(\mathbf{x})$.

The determination of the initial infected population $I_0(\mathbf{x})$ and the parameters of the model will be discussed later. After preparing the mesh, parameters, and initial data, we use the finite element method package in *MATLAB* to numerically approximate the model equations, which gives the values of $S(x,t)$ and $I(x,t)$ at the 23,772 mesh nodes (See Figure 3 for the mesh with 552 nodes).

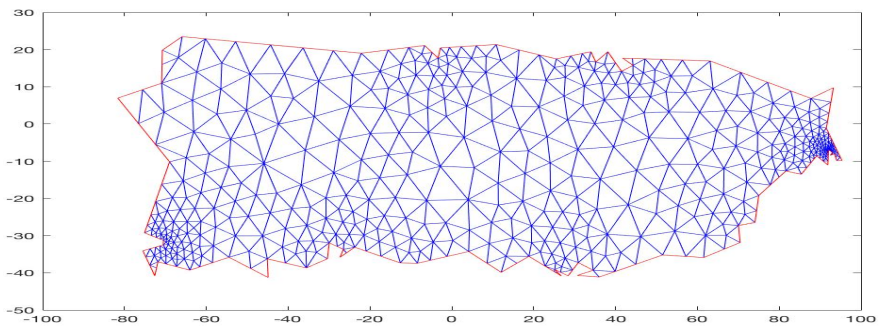

Figure 3: *The geographical mesh with 552 nodes. In the simulations, 23772 nodes are used. The spatial units are kilometers.*

Another difficulty in the simulations is to determine the number of infected cases in a particular municipality. Here, we obtain the administrative boundary data of the 76 municipalities of Puerto Rico from http://gadm.org, which is stored in a *shapefile*. Then, we determine which municipality each small triangle in the mesh belongs to, and compute the number of infected cases in a particular municipality by approximating the integral of $I(x,t)$ on all its triangles (the number of infected cases in a triangle is computed as the average of the values of $I(x,t)$ at its three nodes times its area).

4.1 Parameterization of the Model for Puerto Rico

The parameterization of any model of a seasonal influenza epidemic presents enormous challenges, because of the incompleteness of data. In the United States, typical epidemic data consist of

Morbidity and Mortality Weekly Reports (MMWR) published by the Centers for Disease Control (CDC). For seasonal influenza, these data are very incomplete, and record only a small fraction of total cases. Recent analyses have argued that unreported cases and attack rates (the fractions of the total susceptible populations that become infected over the course of an epidemic) are largely underestimated. In [51] a statistical estimator model was used to estimate the ratio of unreported to reported cases for the H1N1 influenza epidemic in the United States from April to July 2009 as 79-1, and the ratio of total cases to confirmed cases as 140-1.

In an earlier study we developed a formalism for estimating the ratio of reported to unreported cases for the seasonal influenza epidemics in Puerto Rico in 2015-2016 and 2016-2017 [44]. The estimates in [44] claimed attack rates of approximately 40% to 50% with a ratio of unreported to reported cases of approximately 38-1. These attack rates are higher than usually claimed for seasonal influenza epidemics. In the analysis in [44], mass action rather than nonlinear incidence form of disease transmission was used, which may have over-estimated the attack rate.

In [55] it was estimated that approximately 60.8 million cases (attack rate approximately 19%) occurred in the US 2009 H1N1 influenza epidemic, with the US population approximately 325 million during this time period. In [10] data for the 2010-2011 influenza season in the United States was obtained by landline telephone survey of approximately 90,000 people. Of these, 8.9% of adults and 33.9% of children reported influenza like symptoms. In [14] the reporting of serological data is analyzed with respect to measurement errors. In [56], the authors report that 5%-20% of the population in the United States are infected with influenza each year.

Here we have developed our parameters to reflect an attack rate of approximately 18% for the 2016-2017 epidemic in Puerto Rico, with the ratio of unreported to reported cases as approximately 24-1, based on a comparison of the graphs of the reported cases from CDC data and the graphs of the total cases, both reported and unreported, obtained from our model simulations. The objective was to match the duration of the epidemic, the turning point, and the character of the reported cases graph.

Based on these considerations we estimate the parameters for Puerto Rico as follows:

1. Time units are weeks. For the 2016-2017 epidemic, the initial time $t = 0$ corresponds to week 43 of 2016. The 2016-2017 epidemic lasts approximately 18 weeks (http://www.salud.gov.pr/Estadisticas-Registros-y-Publicaciones/Pages/Influenza.aspx).

2. Spatial units are kilometers. The spatial region Ω is as in Figure 1 and 3.

3. The average length of the infectious period of infected people is about 2 days or 1/3.5 weeks: $\lambda(\mathbf{x}) = 3.5\, per\ week$ (http://www.salud.gov.pr/Estadisticas-Registros\-y-Publicaciones/Pages/Influenza.aspx).

4. The transmission parameters for the nonlinear incidence form are $\tau(\mathbf{x}) = 0.02\ per\ week$, $\kappa(\mathbf{x}) = 0.01$, $p = 1.0$, and $q = 1.0$.

5. The diffusion parameter of infected individuals is $\alpha = 2.0\ km^2/week$, which corresponds indirectly to the geographical spread of the virus.

4.2 Simulation of the model for the 2016-2017 epidemic

The model simulates the 2016-2017 seasonal influenza epidemic for all infected cases, not only reported cases. The total reported cases (Departamento de Salud, Puerto Rico) and the total

simulated cases are graphed in Figure 4. The estimated number of total infected cases is 7.532×10^5 with attack rate 18.62% for the seasonal influenza 2016-2017 epidemic. The total cases of the model simulation reflect the total number of reported cases. The ratio of total cases from the model simulation to the total reported cases from Departamento de Salud is 24-1.

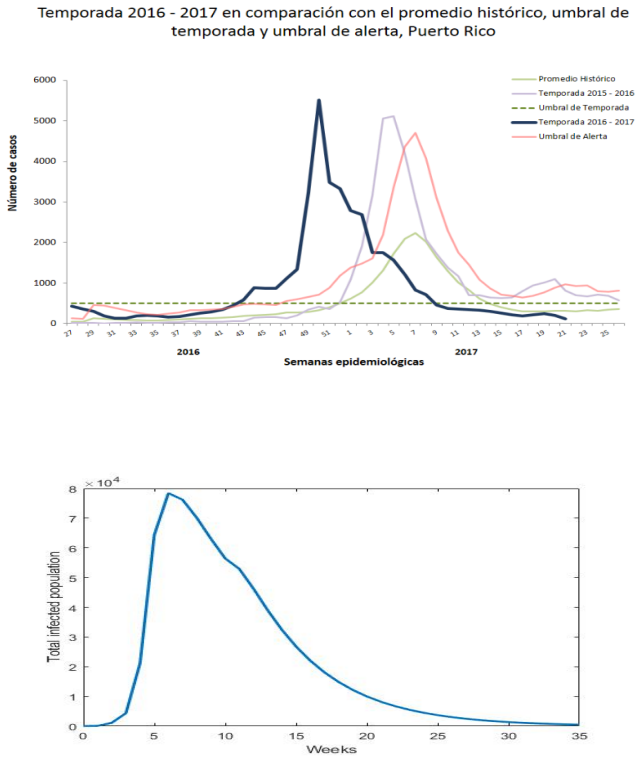

Figure 4: *(top) Reported cases of seasonal influenza Puerto Rico in 2015-2016 (yellow graph) and 2016-2017 (black graph); (bottom) Total cases from the model simulation for 2016-2017.*

For the comparison of the geographical distribution of cases from the model simulation and the geographical distribution of cases reported by Departamento de Salud, we graph in Figure 5 the reported case data of the four major municipalities Mayaqűez, Arecibo, San Juan, and Ponce from data in http://www.salud.gov.pr/Estadisticas-Registros-y-\Publicacione. Although this data is very rough and scattered, we see that the epidemic breaks out near Mayaqűez, then spreads to San Juan and Arecibo, and then to Ponce.

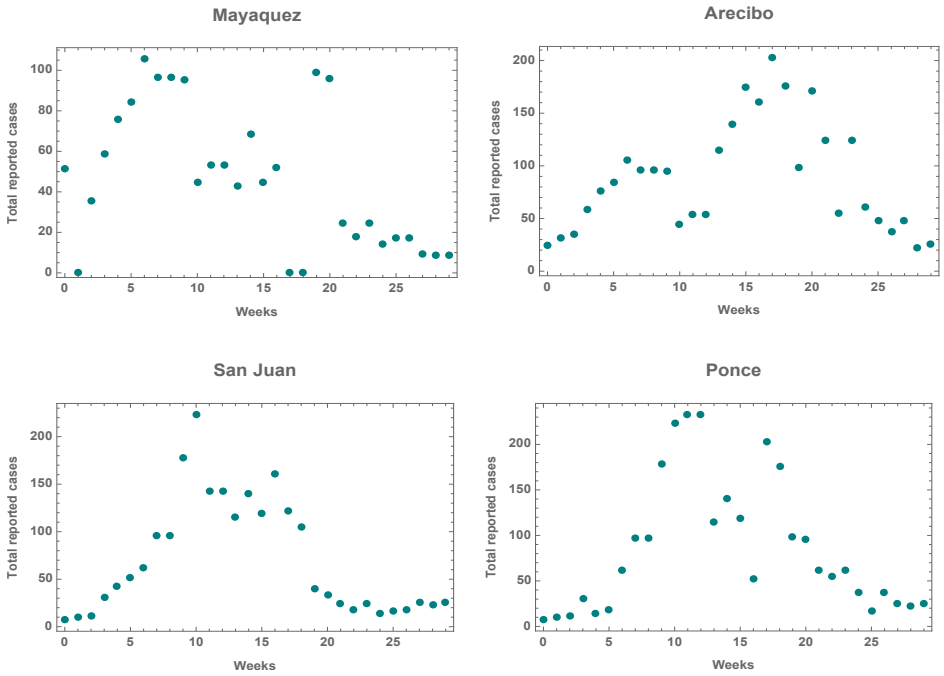

Figure 5: *Estimated reported case data (per 100,000 inhabitants) for four municipalities Mayaqűez, Arecibo, San Juan, and Ponce in the 2016-2017 seasonal influenza epidemic in Puerto Rico. The epidemic arises in Mayaqűez, spreads to Arecibo and San Juan, and last to Ponce.*

For the model simulation, we assume that there are initially, at time 0 (week 43 of 2016), thirty cases normally distributed near Mayaqűez. We then solve the model numerically using the finite element package in *Matlab*. In Figure 6, we graph the simulated total infected cases in the four major municipalities of Puerto Rico. The simulation shows that the epidemic starts in Mayaqűez, spreads to Arecibo and San Juan, and then to Ponce six weeks later, which agrees with the reported case data in Figure 5. We remark that the information in the curves in Figure 5 is meaningful, although the actual values are not accurate. From Figure 6, we can observe that the number of infected cases peaks very quickly once the disease arrives and fades away in about 10 weeks in each municipality, although the epidemic can last much longer in the whole island of Puerto Rico.

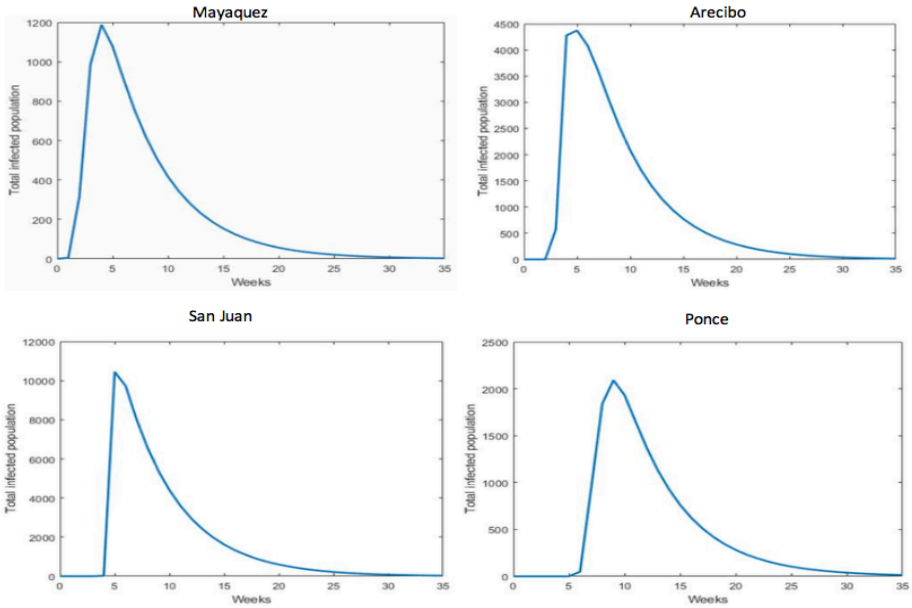

Figure 6: *Model simulation of total cases for four municipalities in the seasonal influenza 2016-2017 epidemic in Puerto Rico.*

We graph the density of the total infected population from the model simulation at weeks 1, 5, 10, 14, 18, 22 in Figure 7, and in all municipalities at weeks 4, 6, 10, and 18 in Figure 8.

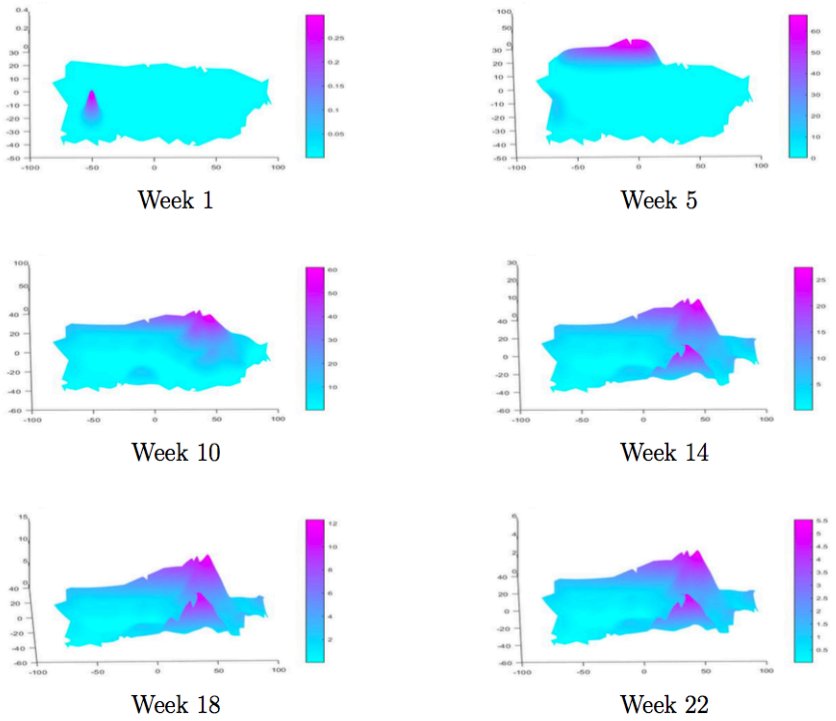

Figure 7: *Simulation of spatial spread of 2016-2017 influenza outbreak in Puerto Rico. The population density of Puerto Rico is set as the initial value of the susceptible population. The initial size of the infected population is assumed to be 30, concentrated in the northwest.*

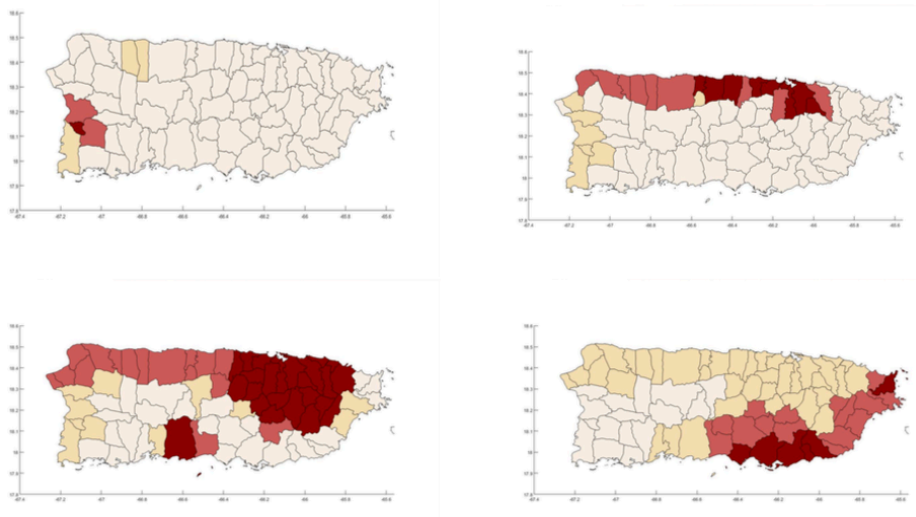

Figure 8: *Model simulation of the infected population densities (number of cases per 100,000 people) in the 2016-2017 seasonal influenza epidemic in Puerto Rico in all municipalities for weeks 4 (top left), 6 (top right), 10 (bottom left), and 18 (bottom right).*

We remark that the reported case data in Figure 4(a) for the 2015-2016 influenza epidemic in Puerto Rico shows two peaks, a main peak in the early stage and a small second peak in the late stage. One possible explanation for the two peaks could be the geographical time evolution of the epidemic, with late arrival in some region. In this case, however, the second peak is apparently due to a second outbreak of a different strain than the strain of the initial outbreak (Figure 9, 10). We also remark that for standard forms of disease transmission in ordinary differential equations epidemic models (such as nonlinear incidence), the infected population can have at most one peak ([41], [42], Theorem 3.1).

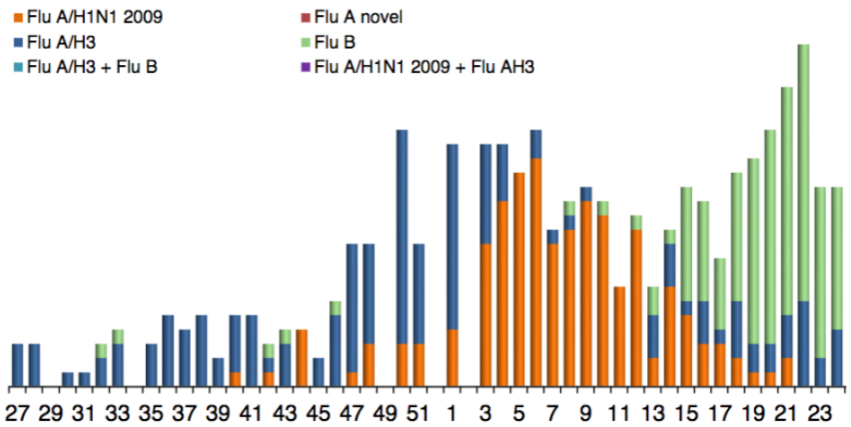

Figure 9: *The total number of reported cases of influenza strain subtypes in 2015-2016. An outbreak of type B strain peaks at week 21 in 2016 (Departamento de Salud, Puerto Rico).*

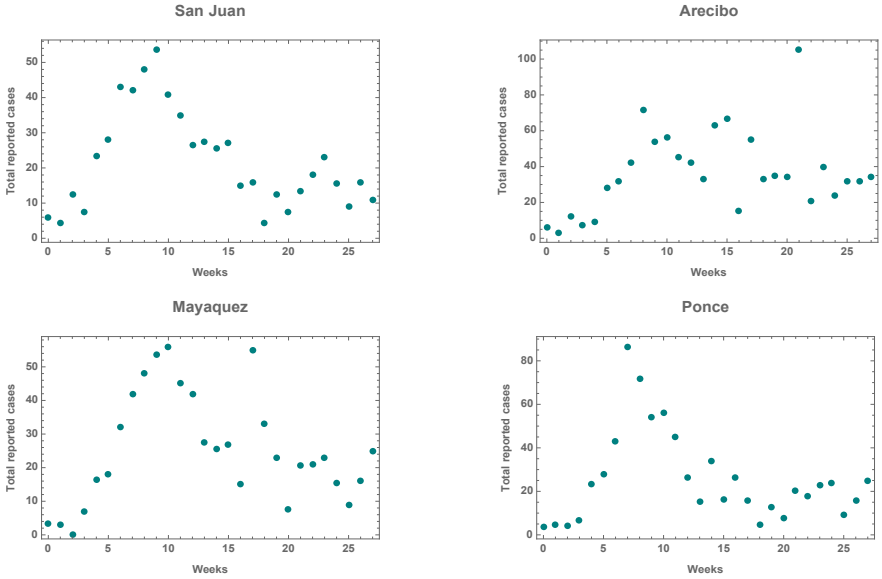

Figure 10: *Estimated reported case data (per 100,000 inhabitants) from Departamento de Salud for four municipalities San Juan, Arecibo, Ponce, and Mayaqűez in the 2015-2016 seasonal influenza epidemic in Puerto Rico. The late second peak is present in all four municipalities.*

5 Conclusions and discussion

Our model indicates that influenza in Puerto Rico rises each season from initial small outbreak locations, and spreads through most of the island, dependent on geographic population variation. The initial outbreak locations and geographic heterogeneity of the population have a large impact on the infected population density over time. In a general region, the epidemic lasts approximately 20-25 weeks, but in subregions the epidemic last approximately 6-12 weeks. The model indicates that the epidemic duration depends strongly on the depletion of the susceptible population to a level that no longer sustains transmission. This depletion happens rapidly in local regions, while the general level of the epidemic occurs much longer in larger regions. Thus, geographic variation is important in understanding the temporal evolution of seasonal influenza epidemics.

The model indicates that the most effective controls are to monitor the importation of infected people into local regions, and to concentrate public health interventions in regions of high population density, especially at the beginning of the epidemic outbreak. Measures such as compulsory quarantine, targeted vaccination, school closings, public event cancellations, and other public health policies can be implemented rapidly in strategic locations, if the severity of the epidemic is recognized in the beginning stages. Future work involves the use of disease age to track infectiousness levels of infected individuals, through an incubation period, and the rise and fall of the infectious period. Particular emphasis will be given to pre-symptomatic infectiousness periods. The model will be extended to include public policy measures such as quarantine, vaccination, and school closings. Future work will extend the model to study geographic variation in other diseases, including vector-borne diseases such as zika, dengue, and malaria.

References

[1] L. J. S. Allen, B. M. Bolker, Y. Lou and A. L. Nevai, Asymptotic profiles of the steady states for an SIS epidemic reaction-diffusion model, *Discrete and Continuous Dynamical Systems*, **21(1)** (2008), 1-20.

[2] N.D. Alikakos, L^p bounds of solutions of reaction-diffusion equations, *Communications in Partial Differential Equations*, **4(8)** (1979), 827-868.

[3] J. Arino, Spatio-temporal spread of infectious pathogens of humans, *Infect. Disease Model*, **2(2)** (2017), 218-228.

[4] J. Arino and K. Khan, Using mathematical modelling to integrate disease surveillance and global air transportation data, in "*Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases*" (eds. D. Chen, B. Moulin, and J. Wu), John Wiley & Sons (2014).

[5] J. Arino and S. Portet, Epidemiological implications of mobility between a large urban centre and smaller satellite cities, *J. Math. Biol.*, **71** (2015), 1243-1265.

[6] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, and A. Vespignani, Multiscale mobility networks and the spatial spreading of infectious diseases, *PNAS USA*, **106(51)** (2009), 21484-21489.

[7] N. F. Britton, An integral for a reaction-diffusion system, *Appl. Math. Lett.* **4** (1991), 43-47

[8] A. Ducrot and T. Giletti, Convergence to a pulsating travelling wave for an epidemic reaction-diffusion system with non-diffusive susceptible population, *J. Math. Biol.* **69** (2014), 533-552

[9] D. Bandaranayake, M. Jacobs, M. Baker, *et al.*, The second wave of 2009 pandemic influenza A (H1N1) in New Zealand, January-October 2010, *Eurosurveillance*, **16(6)** (2011), 1978.

[10] M. Biggerstaff and L. Balluz, Self-reported influenza-like illness during the 2009 H1N1 influenza pandemic, *US Morbid. Mortal. Weekly Rep.*, September 2009 March 2010, **60** (2011), 60:37.

[11] E. Bonabeau, L. Toubiana, and A. Flahault, The geographical spread of influenza, *Proc. Roy. Soc. Lond. B*, **265** (1998), 2421-2425.

[12] N.F. Britton, An integral for a reaction-diffusion system, *Appl. Math. Lett.*, **4(1)** (1991), 43-47.

[13] V. Capasso, Global solution for a diffusive nonlinear deterministic epidemic model, *SIAM J. Appl. Math.* **35(2)** (1978), 274-284.

[14] S. Cauchemez, P. Horby, A. Fox, *et al.*, Influenza infection rates, measurement errors and the interpretation of paired serology, *PLOS Pathog.*, **8(12)** (2012).

[15] S. Charaudeau, P. Khashayar, and P.-Y. Boelle, Commuter mobility and the spread of infectious diseases: application to influenza in France, *PLOS One* **9(1)** (2014).

[16] V. Charu, S. Zeger, J. Gog, *et al.*, Human mobility and the spatial transmission of influenza in the United States, *PLOS Comput. Biol.* **1005382** (2017).

[17] B.J. Coburn, G.W. Bradley and S. Blower, Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1), *BMC Med.* **7(1)** (2009).

[18] V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron and A. Vespignani, Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions, *PLOS Med.*, **4(1)** (2007).

[19] R. Cui, K.-Y. Lam and Y. Lou, Dynamics and asymptotic profiles of steady states of an epidemic model in advective environments, *J. Differential Equations*, **263(4)** (2017), 2343-2373.

[20] A. Ducrot and T. Giletti, Convergence to a pulsating travelling wave for an epidemic reaction-diffusion system with non-diffusive susceptible population, *J. Math. Biol.*, **69** (2014), 533-552.

[21] L. Dung, Global L^∞ Estimates for a Class of Reaction?Diffusion Systems, *Journal of Mathematical Analysis and Applications*, **217**(1998): 72-94.

[22] S. Eubank, H. Guclu, V.S.A. Kumar, and M.V. Marathe, Modelling disease outbreaks in realistic urban social networks, *Nature*, **429** (2004), 180.

[23] N.M. Ferguson, D.A. Cummings, S. Cauchemez and C. Fraser, Strategies for containing an emerging influenza pandemic in Southeast Asia, *Nature*, **437** (2005), 209.

[24] W.E. Fitzgibbon and M. Langlais, Simple models for the transmission of microparasites between host populations living on noncoincident spatial domains, in *Structured population models in biology and epidemiology* (editors. P. Magal and S. Ruan). Springer-Verlag, 2008.

[25] W.E. Fitzgibbon, J.J. Morgan, and G.F. Webb, An outbreak vector-host epidemic model with spatial structure: the 2015-2016 zika outbreak in Rio de Janeiro, *Theor. Biol. Med. Model.*, **14(1)** (2017), 7.

[26] W.E. Fitzgibbon, J.J. Morgan, G.F. Webb, and Y. Wu, A vector-host epidemic model with spatial structure and age of infection, *Nonlinear Analysis: Real World Applications*, **41** (2017), 692-705.

[27] T.C. Germann, K. Kadau, I.M. Longini, and C.A. Macken, Mitigation strategies for pandemic influenza in the United States, *PNAS USA*, **103** (2006), 5935-5940.

[28] J.R. Gog, S. Ballesteros, C. Viboud, L. Simonsen, *et al.*, Spatial transmission of 2009 pandemic influenza in the US, *PLOS Comput. Biol.*, **10(6)** (2014).

[29] R.F. Grais, J.H. Ellis, and G.E. Glass, Assessing the impact of airline travel on the geographic spread of pandemic influenza, *Eur. J. Epidemiol.*, **18(11)** (2003), 1065-1072.

[30] H.W. Hethcote and P. van den Driessche, Some epidemiological models with nonlinear incidence, *J. Math. Biol.*, **29** (1991), 271-287.

[31] L. Hufnagel, D. Brockmann and T. Geisel, Forecast and control of epidemics in a globalized world, *PNAS USA*, **101(42)** (2004), 15124-15129.

[32] A. Huppert, O. Barnea, G. Katriel, *et al.*, Modeling and statistical analysis of the spatio-temporal patterns of seasonal influenza in Israel, *PLOS One*, **7(10)** (2012).

[33] A. Kallen, P. Arcuri, and J. D. Murray, A simple model for the spatial spread and control of rabies, *J. Theor. Biol.*, **116(3)** (1985), 377-393.

[34] I.Z. Kiss, J.C. Miller, and P.L. Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models*, Interdisciplinary Applied Mathematics, **46**, Springer Nature (2017).

[35] T. Kuniya and J. Wang, Lyapunov functions and global stability for a spatially diffusive SIR epidemic model, *Appl. Anal.*, **96(11)** (2017), 1935-1960.

[36] J.P. LaSalle, Some extensions of Liapunov's second method. *IRE Transactions on circuit theory*, **7(4)** (1960), 520-527.

[37] H. Li, R. Peng and F.-B. Wang, Varying total population enhances disease persistence: qualitative analysis on a diffusive SIS epidemic model, *J. Differential Equations*, **262** (2017), 885-913.

[38] W.M. Liu, H.W. Hethcote, and S.A. Levin, Dynamical behavior of epidemiological models with nonlinear incidence rates, *J. Math. Biol.*, **25** (1987), 359-380.

[39] E.T. Lofgren, J.B. Wenger, N.H. Fefferman, *et al.*, Disproportional effects in populations of concern for pandemic influenza: insights from seasonal epidemics in Wisconsin, 1967-2004, *Influenza Other Resp.* **4(4)** (2010), 205-212.

[40] I.M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D.A. Cummings and M.E. Halloran, Containing pandemic influenza at the source, *Science*, **309** (2005), 1083-1087.

[41] P. Magal, O. Seydi, and G.F. Webb, Final size of an epidemic for a two-group SIR model, *SIAM J. Appl. Math.*, **76(5)** (2016), 2042-2059.

[42] P. Magal, O. Seydi, and G.F. Webb, Final size of a multi-group SIR epidemic model: Irreducible and non-irreducible modes of transmission, *Math. Biosci.*, **301** (2018), 59-67.

[43] P. Magal, and H.R. Thieme, Eventual compactness for a semiflow generated by an age-structured models. *Communications on Pure and Applied Analysis*, **3** (2004), 695-727.

[44] P. Magal and G.F. Webb, The parameter identification problem for SIR epidemic models: identifying unreported cases, *J. Math. Biol.* **13**, (2018).

[45] R.H. Martin, *Nonlinear operators and differential equations in Banach spaces.* Wiley, New York (1976).

[46] N. Masuda and P. Holme, *Temporal Network Epidemiology, Theoretical Biology*, Springer Nature Singapore, 2017.

[47] S. Merler and M. Ajelli, The role of population heterogeneity and human mobility in the spread of pandemic influenza, Proc. R. Soc. B **277** (2010), 557-565.

[48] M. Moorthy, D. Castronovo, A. Abraham, *et al.*, Deviations in influenza seasonality: odd coincidence or obscure consequence? *Clin. Microbiol. Infec.*, **18(10)** (2012), 955-962.

[49] J. D. Murray, *Mathematical biology*, 3rd edn. Springer, Berlin (2003).

[50] L. Rass and J. Radcliffe, *Spatial Deterministic Epidemics*, **102**, American Mathematical Society, 2003.

[51] C. Reed, F.J. Angulo, D.L. Swerdlow, M. Lipsitch, M.I. Meltzer, D. Jernigan, and L. Finelli, Estimates of the prevalence of pandemic (H1N1) 2009, United States, April-July 2009, *Emerg. Infect. Dis.*, **15(12)** (2009).

[52] S. Ruan and W. Wang, Dynamical behavior of an epidemic model with a nonlinear incidence rate, *J. Differential Equations*, **1(10)** (2003), 135-163.

[53] S. Ruan, Spatial-temporal dynamics in non local epidemiological models, in *Mathematics for Life Science and Medicine* (eds. Y. Takeuchi, Y. Iwasa, and K. Sato) Springer Berlin Heidelberg, (2007).

[54] L.A. Rvachev and I.M. Longini, A mathematical model for the global spread of influenza, *Math. Biosci.*, **75(1)** (1985), 3-22.

[55] S.S. Shrestha, D.L. Swerdlow, R.H. Borse, *et al.*, Estimating the burden of 2009 pandemic influenza A (H1N1) in the United States (April 2009-April 2010), *Clin. Infect. Dis.,* Jan 1;52 Suppl 1:S75-82, (2011).

[56] K.M. Sullivan, A.S. Monto, and I.M. Longini, Estimates of the US health impact of influenza, *Am. J. Public Health*, **83(12)** (1993), 1712-1716.

[57] N. K. Vaidya, F.-B. Wang and X. Zou, Avian influenza dynamics in wild birds with bird mobility and spatial heterogeneous environment, *DCDS B*, **17(8)** (2012), 2829-2848.

[58] P. Waltman, *Deterministic Threshold Models in the Theory of Epidemics*, Lecture Notes in Mathematics, Springer-Verlag, 1974.

[59] W. Wang and X.-Q. Zhao, Basic reproduction numbers for reaction-diffusion epidemic models, *SIAM J. Appl. Dyn. Syst.*, **11(4)** (2012), 1652-1673.

[60] G.F. Webb, Compactness of Bounded Trajectories of Dynamical Systems in Infinite Dimensional Spaces, *Proc. Roy. Soc. Edinburgh.* **84A**, 19-33 (1979).

[61] G.F. Webb, A reaction-diffusion model for a deterministic diffusive epidemic, *J. Math. Anal. Appl.*, **84(1)** (1981), 150-161.

[62] J.B. Wenger and E.N. Naumova, Seasonal synchronization of influenza in the United States older adult population, *PLOS One*, **5(4)** (2010).

[63] S.M. Zimmer, C.J. Crevar, D.M. Carter, *et. al.*, Seroprevalence following the second wave of Pandemic 2009 H1N1 influenza in Pittsburgh, PA, USA, *PLOS One,* **5(7)** (2010).