A Spatial-Temporal Model for the Evolution of the COVID-19 Pandemic in Spain Including Mobility

Francesc Aràndiga, Antonio Baeza *, Isabel Cordero-Carrión®, Rosa Donat, M. Carmen Martí, Pep Mulet and Dionisio F. Yáñez®

Departament de Matemàtiques, Universitat de València, Av. Vicent Andrés Estellés, E-46100 Burjassot, Spain; francesc.arandiga@uv.es (F.A.); isabel.cordero@uv.es (I.C.-C.); rosa.m.donat@uv.es (R.D.); maria.c.marti@uv.es (M.C.M.); jose.mulet@uv.es (P.M.); dionisio.yanez@uv.es (D.F.Y.)

* Correspondence: antonio.baeza@uv.es

Received: 20 August 2020; Accepted: 23 September 2020; Published: 1 October 2020

Abstract: In this work, a model for the simulation of infectious disease outbreaks including mobility data is presented. The model is based on the SAIR compartmental model and includes mobility data terms that model the flow of people between different regions. The aim of the model is to analyze the influence of mobility on the evolution of a disease after a lockdown period and to study the appearance of small epidemic outbreaks due to the so-called imported cases. We apply the model to the simulation of the COVID-19 in the various areas of Spain, for which the authorities made available mobility data based on the position of cell phones. We also introduce a method for the estimation of incomplete mobility data. Some numerical experiments show the importance of data completion and indicate that the model is able to qualitatively simulate the spread tendencies of small outbreaks. This work was motivated by an open call made to the mathematical community in Spain to help predict the spread of the epidemic.

Keywords: spatial-temporal SAIR model; mobility; outbreak spread; COVID-19; epidemic model; lockdown

1. Introduction

1.1. Motivation And Scope

On 31 December 2019, the World Health Organization (WHO) office in the People’s Republic of China got an statement from the Wuhan Municipal Health Commission reporting cases of ‘viral pneumonia’ in Wuhan. Since then, the disease, known today as COVID-19, caused by the SARS-CoV-2 virus, has become a part of everyone’s daily life by now. The COVID-19 was declared pandemic by the WHO on 11 March 2020.

In Spain, the first case of COVID-19 was diagnosed on 31 January 2020. During February, the number of infected people increased quickly, as well as the number of deaths, and on 14 March 2020 the state of alarm was declared and a lockdown was imposed on all the country as a drastic measure to dramatically reduce the spread of the disease. From that day on, people should stay at home and only indispensable trips within their hometowns were allowed. All flights from or to Spain were canceled, except the ones launched by the government bringing people home. However, some essential activities were allowed as, for instance road transportation of goods or harvesting fruits and vegetables to export them.

These strict mobility measures were lessened gradually. A four-phase de-escalation plan was designed by the government and applied from 2 May 2020 to 22 June 2020, when the state of alarm ended, along with the lockdown and mobility restrictions. Each phase consisted on different contention...
policies. Regarding mobility, the most important changes came with the second phase (Phase 1) that allowed people to travel freely inside their province and with the third phase (Phase 2) that allowed people to travel freely through all the country. It is important to remark that mobility was permitted as long as no symptoms of the infection were shown.

Answering the call from the Comité Español de Matemáticas, (CEMat, Spanish Committee of Mathematics, see http://matematicas.uclm.es/cemat/covid19/en/), it is the purpose of this paper to model the spreading of the COVID-19 virus in Spain, focusing on the impact of the mobility restrictions imposed to the population during the state of alarm. This work is based on previous research on the spread of the 2009 A/H1N1 influenza pandemic in Chile [1].

To estimate the impact of people’s mobility on the evolution and spread of the COVID-19 disease, it is really important to have access to real mobility data, so that models that include mobility data terms with parameters that can be tuned or estimated from said real data can be used. The Spanish National Institute of Statistics (INE) gave access to the research community to some mobility data obtained from the position of cell phones. These data were provided by the three major phone companies in Spain, covering around 80% of the total operating smartphones in the country.

The provided data divide the country in smaller areas, depending on the density of population, and give information about the flux of people moving from their residency area to another one during the state of alarm. These data can be compared with the same fluxes happening on a normal day, before the mobility restrictions, to study how these mobility restrictions imposed really affected the spread of the infection.

In accordance with the mobility data at hand, the model presented in this work divides the spatial domain under consideration into an arbitrary number $C \in \mathbb{N}$ of smaller pieces, called mobility areas, corresponding to the mobility areas defined in the data provided by the INE or to aggregates of them. Moreover, the population will be also divided into groups by means of a compartmental model, that will be defined for each mobility area under consideration.

The mobility model described in this work is strongly influenced by the nature of the mobility data supplied by the authorities. The model can be applied to any other country or region as far as similar mobility fluxes are available or can be reliably estimated.

1.2. Related Work

The most used models for disease transmission are compartmental models. In compartmental models for epidemiology [2–4], a population is divided into groups (called compartments) according to the state in which individuals are regarding the disease. The model is defined by the type of compartments considered and by the statement of the ways in which individuals may change from one compartment to another. Both issues are strongly influenced by the particular disease under consideration, but different models, based on different hypotheses, can be raised to model the same outbreak.

Qualitatively, there exist two major groups of compartmental models: deterministic and stochastic. Stochastic models [5] are based on the fact that some variables or parameters are random variables and the phenomenon being modelled is a stochastic process. Typical techniques used in these models are continuous-time Markov chains and stochastic differential equations. Deterministic models assume that the variables and parameters are continuous functions or constants, and the models are based on ordinary or partial differential equations.

In this work we will focus on deterministic models. The state variables are the number of individuals in each group. One of the simplest model is the SIR model, first introduced in Reference [6]. This model consists on considering three compartments:

- **Susceptible (S)**: individuals that can be infected with the disease,
- **Infectious (I)**: individuals that have the disease and can transmit it,
- **Retired (R)**: individuals that have recovered from the disease, or have died.
Note that this model considers either that the disease does not have an incubation period, or that individuals can transmit the disease during this period. No change in population because of new people born is included in the model.

Individuals change compartment according to some fluxes (units are individuals per unit of time), typically governed by flux rates (percentage of individuals in one group that move to another group per unit of time). Flux rates are normally taken as constants to be determined from empirical data, through some kind of parameter estimation, but in a more realistic scenario non-constant flux rates have to be considered.

If a perfect mixing of individuals in each compartment is assumed, the model can be posed as a system of Ordinary Differential Equations (ODE), where the left hand sides are the variations in each group and the right hand sides are built from fluxes. For instance, the SIR model corresponds to the ODE system

\[
\begin{align*}
S' &= -\beta \frac{IS}{N}, \\
I' &= \beta \frac{IS}{N} - \gamma I, \\
R' &= \gamma I,
\end{align*}
\]

(1)

with total population \(N = S + I + R\) and \(\beta, \gamma\) positive constants. Important quantities in the model are the infection rate \(\beta \frac{I}{N}\) (probability of getting infected in one unit of time) at which susceptible individuals get infected, which is proportional to the proportion of infected individuals \(\frac{I}{N}\) through the contact rate \(\beta\).

The recovery rate \(\gamma\) is equal to \(1/\bar{T}_R\), where \(\bar{T}_R\) is the mean infectious period measured in days. An interesting parameter of the model is the basic reproduction number, denoted by \(R_0\), and given by \(R_0 = \frac{\beta}{\gamma}\). At a very high level the behavior of an epidemic modeled by the SIR model is controlled by this parameter: if \(R_0 > 1\) then the number of (instant) infectious takes its maximum a certain time after the appearance of the epidemic, and then decreases, while if \(R_0 < 1\) then the infectious is always decreasing. In both cases the size of \(I\) tends to zero as time goes to infinity.

A natural extension of the SIR model is the SEIR model [7], that includes Exposed (\(E\) compartment) individuals. This group models the incubation period of the disease, as exposed individuals have been infected but cannot transmit the disease to susceptible individuals. The SEIR model can be written as the following ODE system

\[
\begin{align*}
S' &= -\beta \frac{IS}{N}, \\
E' &= \beta \frac{IS}{N} - \eta E, \\
I' &= \eta E - \gamma I, \\
R' &= \gamma I,
\end{align*}
\]

(2)

with all common symbols having the same meaning as in the SIR model, and \(\eta = 1/\bar{T}_I\), where \(\bar{T}_I\) is the mean incubation period measured in days.

Most attempts made to model the COVID-19 outbreak by means of compartmental models were initially based on the SIR or SEIR models, see for instance References [8–10]. As the epidemic spreads, researchers realized that these models were unable to accurately predict the real evolution of the pandemic, even in individual countries or small regions. As a consequence, most researchers directed their effort towards developing more fitting models, for instance, adding more compartments to the model such as pre-symptomatic, hospitalized, requiring intense care units, quarantined, isolated and exposed, isolated and infected, recovered or dead, see References [11–14], possibly leading to over-parameterized models. Other studies propose age-structured models that divide the population in classes depending on their age [15,16] or coupling them with available mobility data to study the effect of people’s mobility in the spread of the disease [17,18]. This is the approach we will follow in this article.
1.3. Outline of the Paper

The remainder of this paper is organized as follows—T he ingredients of the mathematical model proposed in this work are introduced in Section 2. In Section 2.1 the SAIR compartmental model is introduced as well as all the parameters involved in its definition. In Section 2.2 it is explained in detail how the real mobility data used to build the SAIR model with mobility were collected. The details about how to introduce the mobility restrictions to the SAIR and SIR models with mobility terms are introduced in Sections 2.3 and 2.4, respectively. Section 2.5 deals with the possibility of having time-variable contact rates in the models, and the algorithm used to complete missing information in mobility fluxes is detailed in Section 2.6.

In Section 3 we perform some numerical experiments using real data to validate the proposed model. Finally, in Section 4 some conclusions are drawn.

2. Mathematical Model

2.1. SAIR Model

Our model is based on the less common compartmental model called SAIR (Suspected-Asymptomatic-Infected-Retired) model [19], which substitutes the E compartment of the SEIR model by the A compartment, modelling the infectious asymptomatic individuals. This model is relevant when there are many undetected asymptomatic infectious individuals in the A compartment and symptomatic detected infectious individuals in the I compartment and has been used to model diseases with a very high degree of asymptomatic individuals, as for example dengue [20] and of course COVID-19 [21,22].

Our choice of the SAIR model is motivated by the following facts:

- We do not consider the E group in this work because it is not clear whether people can infect others during the incubation period, nor its duration, hence the value of the parameter $\eta$ in (2) is unknown. A SEAIR model could also be considered if the duration of the non-infectious incubation period is significant.
- According to several studies [23–25], the number of individuals that present little or no symptoms of COVID-19, and therefore cannot be counted as infected, but tests positive, is relevant. In the particular case of Spain, a study conducted by the state health authorities [26] indicates that the number of asymptomatic individuals is really high, reaching a 90% of total infected individuals. Of course, the methods used to detect infections affect these percentages but in any case it seems necessary to consider the A group.
- Including more compartments (e.g., consider individuals hospitalized, or in the intense care unit, split dead individuals from recovered, etc.) would over-parameterize the model and give no relevant information. In addition, the way in which data is provided by the authorities (regional and national) in Spain has changed over time and varies from region to region, making extremely difficult its usage.

For our SAIR model, we suppose that the individuals in the A and I compartments may have different contact rates, denoted by $\beta_A$ and $\beta_I$, and different recovery rates $\gamma_A$ and $\gamma_I$, respectively. The meaning of these quantities is analogous to the SIR and SAIR models, that is, for $X = A, I$, $\frac{\beta_X}{N}$ indicates the probability of getting infected by having contact with an individual of the group $X$ in a unit of time (day) and $\gamma_X = \frac{1}{T_{R,X}}$, where $T_{R,X}$ is the mean infectious period for individuals in the group $X$.

We also consider a rate $\delta$ of individuals in $A$ that may develop symptoms or are anyway detected and hence will move to the $I$ compartment. This parameter depends also on the rapid detection of asymptomatic individuals [27].
Then, the constitutive equations of our SAIR model can be written as:

\[ S' = -\beta_A A + \beta_I I S, \]
\[ A' = \frac{\beta_A A + \beta_I I}{N} S - (\gamma_A + \delta) A, \]
\[ I' = -\gamma_I I + \delta A, \]
\[ R' = \gamma_A A + \gamma_I I, \]

(3)

where now the total population is \( N = S + A + I + R \).

2.2. Mobility Data

In this section we are going to describe in detail the mobility data between provinces and autonomous communities (CCAA) of Spain, provided by the INE, and explain how we used it in our model. The mobility data provided are based on the localization of cell phones. The data cover the three most important phone operators in Spain and they gather more than 80% of the cell phones in the country, so they can be considered realistic.

The study divided the national territory into 3214 mobility areas, which are population groups of between 5000 and 50,000 inhabitants (more homogeneous than the municipalities) and that can have totally different sizes. For instance, in rural areas where the population density is very low, a mobility area will be constituted by several municipalities, up to reach at least 5000 inhabitants, while deeply populated areas, like big cities as Madrid, are subdivided in several mobility areas of about 20,000 inhabitants on average.

From the start of the lockdown period (16 March 2020) up to the end of the state of alarm (22 June 2020), the information for each mobility area is provided, based on the analysis of 24 h periods. For the sake of comparison, mobility data for a reference day, that is, a typical working day before lockdown, was also provided as the average of the period 18–21 November 2019 (Monday to Thursday).

The data provided for each mobility area are:

1. The registered resident population (according to official records).
2. Number of individuals that stay at their area of residence.
3. Number of individuals that leave their area of residence, and number of cells receiving individuals from that cell.
4. The destination mobility areas to which those who leave their area go and the estimated number of people that move to them, with the restrictions indicated below due to privacy issues.

As indicated in Section 1, these mobility data are based on the localization of cell phones. They are based on the following assumptions:

- A person is considered to reside in an area if their phone stays in said area from 0 a.m. to 6 a.m.
- The destination area of that person, is the area where their mobile phone stays most of the time from 10 a.m. to 4 p.m., with a minimum of two hours. Hence, the destination area of a person can be their residence area, any other area, or none (if it does not stay more than two hours in the same area in the period considered).
- The data is normalized with respect to the total population.
- The phone operators that provided the data did not report flows of people smaller than 10 or 15 phones (depending on the phone company) in order to ensure privacy. Moreover, the estimated flow of people moving between each pair of areas is provided only if it is higher than 100 people. Smaller flows are not provided.

The last point raises the problem that there are mobility areas without information about where people have moved to. That means that if only these data are considered, these areas would appear as
isolated areas, although we know that people have left and entered their residence area. Furthermore, if we merge some mobility areas to create a bigger area (such as a province or region) it might result in a loss of data, since one could have a bigger area with an estimated mobility much smaller than its real one.

With the aim of correcting this, and recovering the information of mobility between areas with an estimated mobility flux smaller than 100 people, we have proposed an algorithm that is explained in Section 2.6.

### 2.3. SAIR Model with Mobility

In this section we introduce the model proposed in this work for the analysis of the spread of infected individuals under mobility constraints. For this model, we divide the whole domain area into an arbitrary number of smaller spatial areas. In each spatial area, we use the SAIR model (3), and we add some terms to simulate the movement of people between regions.

Let us consider \( C \in \mathbb{N} \) small spatial areas covering the whole original domain, and define variables \( S_i, A_i, I_i, R_i \) for each area \( i = 1, \ldots, C \), and mobility fluxes \( m_{jk} \) of individuals per day (which is taken as the unit of time) that move from area \( j \) to area \( k \).

We assume that individuals in the \( I \) compartment do not move and that the mobility flux is a daily flux, that is, it is positive half a day and negative the other half, representing, for example, people going to work and back. This means that, at the first half of the day there are \( m_{jk} \) individuals that move from area \( j \) to area \( k \), i.e, from the compartments \( S_j, A_j, R_j \) to the corresponding \( S_k, A_k, R_k \), while in the second half of the day the same amount of people go back to area \( j \), but some of them have got infected during their stay at area \( k \). If \( \mu_k \) denotes the probability of infection for individuals in the \( k \)-th compartment per day and taking into account that \( m_{jk} \) susceptible individuals have traveled to area \( k \) and have been exposed to the disease for half a day in area \( k \), then the number of individuals that got infected in the \( k \)-th compartment is given by

\[
\frac{1}{2} \mu_k m_{jk} \frac{S_j}{S_j + A_j + R_j}.
\]  

The factor \( \frac{1}{2} \) has been introduced because the time unit is a complete day and individuals stay just half a day in area \( k \).

According to the SAIR model, the probability of getting infected for a susceptible individual in area \( j \) is given by

\[
\mu_j = \frac{\beta_{A,j} A_j + \beta_{I,j} I_j}{N_j},
\]

with \( N_j = S_j + A_j + I_j + R_j \). Hence, for each area \( j \), the SAIR model without mobility (3) is given by

\[
S_j' = -\mu_j S_j,
A_j' = \mu_j S_j - (\gamma_A + \delta) A_j,
I_j' = -\gamma_I I_j + \delta A_j,
R_j' = \gamma_A A_j + \gamma_I I_j.
\]

Note that different contact rates \( \beta_{A,j} \) and \( \beta_{I,j} \) have been considered for different areas, while the same values for \( \gamma_A, \gamma_I \) and \( \delta \) are taken for all areas. This is motivated by the fact that the latter values are related to the evolution of the disease itself while the former depend on the behavior of the susceptible individuals, that may vary from area to area. Although there is no difficulty on considering different values for \( \delta \) and/or \( \gamma \) in each area it would result on an unnecessary over-parameterization.

The susceptible individuals leaving area \( j \) in the first half of the day is given by
\[
\sum_{k} m_{jk} \frac{S_j}{N_j^*},
\]
where \(N_j^* = S_j + A_j + R_j = N_j - I_j\) is the population in the \(j\)-th compartment that might move, while the ones entering area \(j\) in the same period are

\[
\sum_{k} m_{kj} \frac{S_k}{N_k^*}.
\]

Hence, the flux of susceptible individuals in area \(j\) for the first half of the day, normalized for a unit time, is given by the term:

\[
\varphi_1(t) \sum_{k} \left( m_{jk} \frac{S_j}{N_j^*} - m_{kj} \frac{S_k}{N_k^*} \right),
\]

with the function \(\varphi_1(t)\) defined by

\[
\varphi_1(t) = \begin{cases} 
-2, & t - \lfloor t \rfloor \in [0, 1/2), \\
0, & t - \lfloor t \rfloor \in [1/2, 1), 
\end{cases}
\]

where the notation \(\lfloor t \rfloor\) stands for the floor function that gives the greatest integer smaller than or equal to \(t\).

Still regarding the susceptible individuals, for the second half of the day we consider a similar term but subtracting the individuals that got infected during their stay at another area \(k\), given by (4), that move to \(A_k\). Also, some individuals that came from area \(k\) to area \(j\) would move to \(A_j\). This results on a term

\[
\varphi_2(t) \sum_{k} \left( m_{jk} \left( \frac{S_j}{N_j^*} - \frac{1}{2} \frac{S_j}{N_j^*} \mu_k \right) - m_{kj} \left( \frac{S_k}{N_k^*} - \frac{1}{2} \frac{S_k}{N_k^*} \mu_j \right) \right),
\]

with

\[
\varphi_2(t) = \begin{cases} 
0, & t - \lfloor t \rfloor \in [0, 1/2), \\
2, & t - \lfloor t \rfloor \in [1/2, 1), 
\end{cases}
\]

Therefore the evolution equation for susceptible individuals in compartment \(j\) gets the form

\[
S'_j = -\mu_j S_j + \varphi_1(t) \sum_{k} \left( m_{jk} \frac{S_j}{N_j^*} - m_{kj} \frac{S_k}{N_k^*} \right) + \varphi_2(t) \sum_{k} \left( m_{jk} \left( \frac{S_j}{N_j^*} - \frac{1}{2} \frac{S_j}{N_j^*} \mu_k \right) - m_{kj} \left( \frac{S_k}{N_k^*} - \frac{1}{2} \frac{S_k}{N_k^*} \mu_j \right) \right),
\]

By a similar argument the evolution equation for the \(A_j\) compartments is given by

\[
A'_j = \mu_j S_j - (\gamma_A + \delta) A_j + \varphi_1(t) \sum_{k} \left( m_{jk} \frac{A_j}{N_j^*} - m_{kj} \frac{A_k}{N_k^*} \right) + \varphi_2(t) \sum_{k} \left( m_{jk} \left( \frac{A_j}{N_j^*} + \frac{1}{2} \frac{S_j}{N_j^*} \mu_k \right) - m_{kj} \left( \frac{A_k}{N_k^*} + \frac{1}{2} \frac{S_k}{N_k^*} \mu_j \right) \right),
\]

and for the \(R_j\) compartment by

\[
R'_j = \gamma_A A_j + \gamma_I I_j + \varphi_1(t) \sum_{k} \left( m_{jk} \frac{R_j}{N_j^*} - m_{kj} \frac{R_k}{N_k^*} \right) + \varphi_2(t) \sum_{k} \left( m_{jk} \frac{R_j}{N_j^*} - m_{kj} \frac{R_k}{N_k^*} \right),
\]
as individuals in this compartment cannot get infected nor infect others. Finally, the equation for $I_j$ does not depend on mobility as infected individuals do not travel by hypothesis.

Grouping terms the model can be written as a system of 4 C ODEs given by:

$$
S'_j = -\mu_j S_j + \varphi(t) \sum_k \left( m_{jk} \frac{S_j}{N_j} - m_{kj} \frac{S_k}{N_k} \right),
$$

$$
A'_j = \mu_j S_j - (\gamma_A + \delta) A_j + \varphi(t) \sum_k \left( m_{jk} \frac{A_j}{N_j} - m_{kj} \frac{A_k}{N_k} \right),
$$

$$
I'_j = -\gamma I_j + \delta A_j,
$$

$$
R'_j = \gamma_A A_j + \gamma I_j + \varphi(t) \sum_k \left( m_{jk} \frac{R_j}{N_j} - m_{kj} \frac{R_k}{N_k} \right),
$$

(5)

where $\varphi = \varphi_1 + \varphi_2$.

With this setup the system has $C + 3$ parameters. It follows that there are no permanent transfers of population between areas, since the addition of the four equations in (5) yields for each $j = 1, \ldots, C$:

$$
N'_j(t) = \varphi(t) \sum_k (m_{jk} - m_{kj}),
$$

whose solution is

$$
N_j(T) = N_j(0) + \sum_k (m_{jk} - m_{kj}) \int_0^T \varphi(t) \, dt.
$$

Therefore $N_j(T) = N_j(0)$ whenever $T \in \mathbb{N}$, since $\varphi$ is 1-periodic and $\int_0^1 \varphi(t) \, dt = 0$.

Also, note that according to the model, for $n \in \mathbb{N}$

$$
N_j(n + \frac{1}{2}) = N_j(n) + \sum_k (m_{jk} - m_{kj}) \int_n^{n+\frac{1}{2}} \varphi(t) \, dt = N_j(0) + \sum_k (m_{kj} - m_{jk}),
$$

that is, the model predicts that the individuals are one half of the day at their residence area, and the other half at their destination area.

2.4. SIR Model with Mobility

For the sake of comparison, we also consider the SIR model with mobility. In this case asymptomatic and non-asymptomatic infected individuals belong to the same group $I_j$ and they all are infectious. We assume that only a fraction $\epsilon$ of the individuals in $I_j$ (e.g., the asymptomatic ones, as individuals with symptoms will feel sick and stay at home) move from their area. The model is constructed in the same way as the SAIR model with mobility, taking into account that the compartments $A_j$ and $I_j$ are merged together and that the terms describing mobility for the infectious individuals are multiplied by $\epsilon$. The model can be written as a system of 3 C ODE:
various measures have been taken in different phases of the pandemic evolution, to try to dramatically reduce its spread first by a severe confinement on 15 March 2020, and later to recover some economic activity, by means of a 4-phase de-escalation plan, started on 2 May 2020 and finished on 22 June 2020. Each phase consisted on different contention policies.

To model the different stages of the lockdown and de-escalation period regarding mobility restrictions and their effect on the contact rates of the infection, we propose to consider discontinuous functions for the contact rates:

\[
\beta_j(t) = \begin{cases} 
\beta_{j,1} & t \leq t_{j,1} \\
\beta_{j,2} & t_{j,1} < t \leq t_{j,2}, 0 < t_{j,1} < \ldots,
\end{cases}
\]

with \(\beta_{j,1}, \beta_{j,2}, \ldots\) given positive numbers and the times \(t_{j,1}, t_{j,2}, \ldots\) chosen appropriately. The same can be done for the parameters \(\beta_{A,j}\) and \(\beta_{I,j}\) in the SAIR model.

To account for changes in the mobility patterns, there is the possibility of changing the mobility terms \(\phi_i(t)m_{j,k}, i = 1, 2\), by \(\phi_i(t)m_{j,k,l}\), where \(m_{j,k,l}\) is the flux between cells \(j\) and \(k\) on the \(l\)-th day.

Finally, taking into consideration the advances in the medical treatment of the disease, a time-varying function \(\gamma(t)\) could also be considered to model different recovery rates throughout the pandemic evolution in the same fashion we have done with the function \(\beta(t)\). This can also be extended for the parameters \(\gamma_A\) and \(\gamma_I\) in the SAIR model, but it is beyond the scope of this paper.

Although these functions would yield a discontinuous right hand for the system of ODEs, the theory of discontinuous ODEs covers these cases [28], so that there is a unique continuous solution, with discontinuous lateral derivatives at the jumps of the coefficients.
2.6. Completing Fluxes

As mentioned in Section 2.2, in order to estimate the possible missing data when merging mobility areas, we aim to use the information provided as, for instance, the daily sum of all the outgoing fluxes and the number of receiving cells from each source cell. Let \( N_j \) be the population of cell \( j \), \( d_{jk} \) the distance between the centroids of the cells \( j \) and \( k \), and \( g_{jk} = \frac{N_j}{\max(d_{jk}, \epsilon')} \), for some \( \epsilon' > 0 \) (which we take as \( 10^{-5} \) in our experiments). The coordinates of the centroids are the average of the coordinates of the towns in the cell weighted by their population. For the source cell \( j \), we denote by \( n_j \) the number of cells receiving individuals from cell \( j \) and by \( f_j \) the number of individuals leaving cell \( j \) on a given date. The provided fluxes, obtained from real data, between cells \( j \) and \( k \) are denoted by \( m_{jk}^* \). We recall that, as indicated in Section 2.2, these are the fluxes between areas that are higher than 100 people.

Our goal is to define fluxes \( m_{jk} \) between cells \( j \) and \( k \) such that:

1. If \( m_{jk}^* \neq 0 \Rightarrow m_{jk} = m_{jk}^* \) (\( m_{jk} \) “complete” the provided fluxes)
2. \(|\{k/m_{jk} \neq 0\}| = n_j, \sum_k m_{jk} = f_j \) (match the global information of outgoing fluxes).

The strategy is to assign the non-assigned \( \tilde{f}_j = f_j - \sum_k m_{jk}^* \) individuals issuing from cell \( j \) to the first \( \tilde{n}_j = n_j - |\{k/m_{jk}^* \neq 0\}| \) cells in a descending ranking according to \( g_{jk} \):

\[
m_{jk} = \begin{cases} \tilde{f}_j \frac{g_{jk}}{\sum_{l=1}^{\tilde{n}_j} g_{jl}}, & \text{if } k = k_j, l = 1, \ldots, \tilde{n}_j, \\ 0, & \text{otherwise.} \end{cases}
\]

3. Numerical Experiments and Validation

In this section we validate the model presented in Section 2.3 in various scenarios. In Section 3.1 we study how the mobility data and the way in which missing data is completed affects the simulation results. In Section 3.2 we try to simulate a real outbreak corresponding to the region of A Mariña in Spain.

3.1. How Mobility Affects the Spread of Covid-19

In this section we analyze various types of mobility data. In particular three types of mobility are considered:

- Very restricted mobility: we define the mobility fluxes using the INE data from April 1.
- Restricted mobility: the mobility fluxes are obtained with the algorithm in Section 2.6 from the INE data of April 1.
- Unrestricted mobility: in this case, the mobility fluxes are obtained with the algorithm in Section 2.6 but from the INE November reference data.

In Figure 1 the mobility flow matrices for Spain (first row) and for the autonomous communities of Galicia and Valencia, second and third row respectively, can be seen. In the first column we have the movements based on the INE data of April 1 (very restricted mobility) while in the second and third columns the results obtained with the algorithm in Section 2.6 using the INE data of 1 April and November (restricted mobility and unrestricted mobility) respectively, are shown. We only show the movements between areas when there are more than 10 people moving between them, although in the experiments we consider all the movements obtained. In the top row the red squares correspond to autonomous regions, while in the second row they correspond to the provinces in the Valencia region (Alacant, Castelló and Valencia) and in the last row they correspond to the provinces in Galicia (A Coruña, Lugo, Ourense and Pontevedra).
Figure 1. From top to bottom: movements between the 3214 mobility areas organized alphabetically by Autonomous Communities, between the areas of the provinces of Valencia and between the areas of the provinces of Galicia. From left to right: movements between areas of more than 100 people (Spanish National Institute of Statistics (INE) data 1 April), movements between areas obtained with our algorithm from the INE data of April 1 and from the INE data of November. In the two last cases the fluxes are only shown when more than 10 people have moved between areas.

Now we consider some experiments in autonomous communities that we suppose completely isolated from the outside. In these experiments we have set the parameters $\gamma_A = \gamma_I = 1/15$, constant in all areas, $\delta = 1/50$ and $\beta = 3\gamma$. These values have been set according to estimations of the incubation and infectious periods of the disease [29,30], with a slightly high incubation period to make the effect more evident and stand for the time required for the incorporation of individuals to official records. The value of $\beta$ corresponds to a basic reproduction number $R_0 = 3$. For the numerical solution of the model we have used the MATLAB ODE package.

First we consider the Valencian Autonomous Community. We assume that Alicante has 40 asymptomatic infectious individuals and the other two provinces have none. We see in Figure 2 how the development of COVID-19 is different when the different types of mobility are considered.

As a second example, we consider the Autonomous Community of Galicia, where we separate A Mariña from the rest of Lugo province. We assume that A Mariña has 40 asymptomatic infectious individuals and the rest of the provinces have none. In Figure 3 we see the results obtained.
In both Figures 2 and 3 it can be seen that COVID-19 spreads more quickly to closer areas, and that if mobility is less restricted, it appears earlier. In Figure 3 it can be seen that when using very restricted mobility fluxes, COVID-19 does not reach the rest of Lugo province and it is because with these mobility fluxes it is completely isolated from A Mariña and the rest of the provinces.

In these experiments we are considering mobility restrictions between regions or provinces but not within them. This is why the behavior in the regions where infected individuals are initially located (Alacant or A Mariña) is the same independently of the mobility restrictions considered. A finer look to the influence inside each region could be done by considering all mobility areas of the Autonomous Communities (217 in Galicia, 324 in Valencia), but information regarding the number of infected individuals was not given with such detail by the INE. Also, variable parameters in the SAIR model, depending on mobility restrictions, can be considered, as has been done in the experiment of Section 3.2.

**Figure 2.** Simulation of the evolution of COVID-19 in the different provinces of the Valencian Autonomous Community, assuming that initially Alacant has 40 asymptomatic infectious individuals and Castelló and Valencia have none, and considering different mobility restrictions between provinces. From left to right: very restricted mobility, restricted mobility and unrestricted mobility.
Figure 3. Simulation of the evolution of COVID-19 in five regions of the Autonomous Community of Galicia assuming that initially A Mariña has 40 asymptomatic infectious individuals and A Coruña, rest of Lugo (Lugo without A Mariña), Ourense and Pontevedra have none, and considering different mobility restrictions among the five regions. From left to right: very restricted mobility, restricted mobility and unrestricted mobility.

3.2. Outbreak in the A Mariña Area, Galicia, Spain

A Mariña is an area located in the northern part of the province of Lugo, in the autonomous region of Galicia, at the northwest of Spain. It has an extension of about 1650 Km² and a population of
nearly 70,000 people. On 6 July 2020, an outbreak of COVID-19 was declared in that area and mobility restrictions as well as other control measures were dictated so that trips within the area were allowed but limitations were imposed for trips from or to the outside of the area. The lockdown measures were progressively retired, depending on the particular village, up to 22 July 2020, when mobility was again free. Let us summarize the timeline of the outbreak:

- 22 June 2020: End of the state of alarm in Spain. Mobility restrictions do not apply anymore.
- 23 June 2020: A new case is detected in the village of Xove, in the Galician region, A Mariña area.
- 5 July 2020: 106 new cases had been detected in the area since 23 June 2020.
- 6 July 2020: Lockdown is dictated for the area of A Mariña. Restrictions apply for the municipalities of Alfoz, Barreiros, Burela, Cervo, Foz, Lourenzá, Mondoñedo, Ourense, O Vicedo, Viveiro and Xove.
- 12 July 2020: Mobility restrictions were retired for some municipalities.
- 22 July 2020: Mobility restrictions were retired for the full area of A Mariña.
- 20 August 2020: The regional government declared that the A Mariña outbreak is controlled.

Figure 4 shows the official active cases as reported by the authorities [31]. The outbreak in the A Mariña area achieved its maximum on 11 July 2020 and spread to the rest of Galicia. The incidence outside A Mariña due to the spread of the outbreak cannot be split from the rest of the cases, due to other imported cases or to local contacts but the effect of the outbreak in other territories seems evident.

In this subsection we test to what extent our model can simulate such an outbreak. Let us clarify that we do not intend to exactly predict the outbreak figures but to analyze the effects of the various parts of the model and to what extent they can qualitatively represent the outbreak tendencies. This is due to the fact that the data provided by the authorities do not accurately represent the real incidence of the epidemic, because of many factors: undetected cases, unknown number of asymptomatic individuals, different criteria for case reporting among the different autonomous regions, and so forth.

Figure 5 shows the result of various simulations regarding this outbreak. The Galician region has been split into five parts: A Mariña, the complete provinces of Coruña, Ourense and Pontevedra, and the rest of the province of Lugo. The plots show infected individuals, counting for the sum of infected and asymptomatic in the case of the SAIR model. The results have been obtained using the initial data given by 1000 asymptomatic individuals and 100 infected in A Mariña and zero at the rest, in order to take apart the effect of the outbreak from other effects. These data come from the fact that, according to Reference [26], the prevalence of COVID-19 in Spain is about 10 times higher than the
reported cases and on the other hand there is a delay in the process of including data in the official records. In any case, as mentioned, the data does not intend to represent real figures, but tendencies. Let us also mention that the rescaling of the initial data produces results that are not in the same scale as the real reported data. The results have thus been rescaled back so that the numbers shown in Figure 5 are put on a scale similar to the ones in Figure 4. We have set the parameters $\delta = 1/25$, $\gamma_{A,I} = 1/15$, based on estimations of the incubation and infectious periods of the disease [29,30]. The values for the contact rates have been taken as $\beta_A = 5\gamma$ and $\beta_I = 6\gamma$ in all regions. These correspond to basic reproduction numbers $R_{0,A} = \frac{\beta_A}{\gamma_A} = 5$ and $R_{0,I} = \frac{\beta_I}{\gamma_I} = 6$, which may be valid numbers only on the onset of the epidemics. Our present aim is to assess the capability of the model of triggering an outbreak in a location without initial infectious individuals.

**Figure 5.** Predicted active cases (A+I) for different simulations of the A Mariña outbreak. From left to right and from top to bottom: SAIR, with mobility restrictions and reduced contact rates applied on days 14–30; Same experiment with constant contact rates; SAIR, without mobility restrictions and reduced contact rates applied on days 14–30; SIR with restricted mobility on days 14–30; SIR with unrestricted mobility on days 14–30.
In the cases where a reduction of the contact rates is indicated, the above values are reduced by a factor of 2, only in the area of A Mariña. The mobility fluxes have always been completed according to the algorithm described in Section 2.6.

The top left plot in Figure 5 tries to reproduce the real situation in which a lockdown is imposed from 6 July 2020 to 22 July 2020 (days 14–30). The contact rates \( \beta_A \) and \( \beta_I \) have been also reduced in that period, representing the mild official recommendation (not mandatory) of leaving one’s house only if necessary. The figure shows that the outbreak effectively spreads to the rest of the areas. The model shows a peak value of infections at the beginning of August that does not appear in the real data, but this can be due to the rescaling of the data, to the influence of factors different from the outbreak itself or to a limitation of the model.

The top right plot in Figure 5 corresponds to the situation where the contact rates \( \beta_A \) and \( \beta_I \) are not reduced during the lockdown period. As can be seen, the results are very similar, indicating that in this case a mild reduction on the social contact does not have a big impact on the evolution of the outbreak.

The second row in Figure 5 corresponds to the case where no mobility restrictions are applied. In these plots the mobility fluxes have been set to the ones of a normal day before lockdown taken as reference. It can be seen that either with a reduction on the contact rates during days 14–30 (left plot) or without it (right plot) the outbreak spreads faster to the rest of the territories than if restrictions are applied. Also, the peak values of infections are slightly higher.

For comparison, in the last row we show a result obtained with the SIR model including mobility. The left plot corresponds to restricted mobility fluxes, and the right plot to unrestricted fluxes, with constant contact rates. The value \( \varepsilon \) that appears in the model has been set to 0.9, being all other values set to the same values as in the SAIR model. It can be observed that the model also predicts the spread of the disease to the neighboring regions in a way similar to the SAIR model, but with the need of tuning the parameter \( \varepsilon \). Note that in this case the maximum values for infections are slightly smaller, due to a different parameter tuning.

4. Conclusions

In this paper, we present a mathematical model for the simulation of the spread of small epidemic outbreaks. The model is based on the division of a region into smaller areas, called mobility areas, for which mobility fluxes are known, and describe the evolution of an outbreak in one of the pieces by a model which combines a deterministic SAIR compartmental model and terms that include the effects of people’s mobility among areas. The model includes also limited capabilities for the modeling of social confinement measures. As mobility data are sometimes incomplete, we also describe a method for the completion and homogenization of fluxes. Numerical results show that the model is able to qualitatively represent the tendencies of small outbreaks.

**Author Contributions:** Data curation, F.A. and P.M.; funding acquisition, P.M.; investigation, F.A., A.B., I.C.-C., R.D., M.C.M., P.M. and D.F.Y.; methodology, A.B. and P.M.; software, F.A., A.B. and P.M.; supervision, A.B.; Visualization, F.A.; writing—original draft, F.A., A.B., I.C.-C. and P.M.; writing—review & editing, A.B., M.C.M. and D.F.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Spanish MINECO project MTM2017-83942-P, by Spanish Agencia Estatal de Investigación Grant No. PGC2018-095984-B-I0, by Generalitat Valenciana Grant. No. PROMETEO/2019/071 and by European Cooperation in Science and Technology (COST) Grant. No. COST Action CA17137.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.
Nomenclature (Units Are Indicated in Brackets)

Latin symbols

- $A$ : Asymptomatic individuals [individuals]
- $A_j$ : Asymptomatic individuals of area $j$ [individuals]
- $d_{jk}$ : Distance between the centroids of the $j$-th and $k$-th cells [km]
- $E$ : Exposed individuals [individuals]
- $f_j$ : Individuals leaving cell $j$ on a given day [individuals]
- $g_{jk}$ : Quantity used in the algorithm for flux completion [individuals · km$^{-1}$]
- $I$ : Infectious individuals [individuals]
- $I_j$ : Infectious individuals of area $j$ [individuals]
- $m_{jk}$ : Mobility flux from area $j$ to area $k$ [individuals · day$^{-1}$]
- $m^{r}_{jk}$ : Mobility flux from area $j$ to area $k$ provided by real data [individuals · day$^{-1}$]
- $m_{jk,l}$ : Mobility flux from area $j$ to area $k$ on the $l$-th day [individuals · day$^{-1}$]
- $n_j$ : Number of cells that receive individuals from cell $j$ [cells]
- $N$ : Total population [individuals]
- $N_j$ : Total population of area $j$ [individuals]
- $N^*$ : Population in the $j$-th compartment that might move [individuals]
- $R$ : Retired individuals [individuals]
- $R_j$ : Retired individuals of area $j$ [individuals]
- $R_0$ : Basic reproduction number [-]
- $S$ : Susceptible individuals [individuals]
- $S_j$ : Susceptible individuals of area $j$ [individuals]
- $T_I$ : Mean incubation period [days]
- $T_R$ : Mean infectious period [days]
- $T_{R,j}$ : Mean infectious period for individuals in the $j$ compartment [days]

Greek symbols

- $\beta_j$ : Contact rate of compartment $j$ [days$^{-1}$]
- $\beta_{jk}$ : Contact rate of compartment $j$ and area $k$ [days$^{-1}$]
- $\gamma_j$ : Recovery rate of compartment $j$ [days$^{-1}$]
- $\gamma_{jk}$ : Recovery rate of compartment $j$ and area $k$ [days$^{-1}$]
- $\delta$ : Rate of asymptomatic individuals that may develop symptoms [day$^{-1}$]
- $\epsilon$ : Fraction of individuals in the $I$ compartment of the SIR model moving from their area [-]
- $\eta$ : Incubation rate [days$^{-1}$]
- $\mu_j$ : Probability of getting infected for a susceptible individual in area $j$ [-]
- $\varphi_1(t)$ : Auxiliary function used to define the flux of individuals in area $j$ for the first half of the day [-]
- $\varphi_2(t)$ : Auxiliary function used to define the flux of individuals in area $j$ for the second half of the day [-]

References


