



---

*Research article*

## **Geographical network model for COVID-19 spread among dynamic epidemic regions**

**Roman Zúñiga Macías, Humberto Gutiérrez-Pulido, Edgar Alejandro Guerrero Arroyo and Abel Palafox González\***

Universidad de Guadalajara, CUCEI, Blvd. Marcelino García Barragán 1421, 44430, Guadalajara, Jal., México

\* **Correspondence:** Email: [abel.palafox@academicos.udg.mx](mailto:abel.palafox@academicos.udg.mx).

**Abstract:** Pandemic due to SARS-CoV-2 (COVID-19) has affected to world in several aspects: high number of confirmed cases, high number of deaths, low economic growth, among others. Understanding of spatio-temporal dynamics of the virus is helpful and necessary for decision making, for instance to decide where, whether and how, non-pharmaceutical intervention policies are to be applied. This point has not been properly addressed in literature since typical strategies do not consider marked differences on the epidemic spread across country or large territory. Those strategies assume similarities and apply similar interventions instead. This work is focused on posing a methodology where spatio-temporal epidemic dynamics is captured by means of dividing a territory in time-varying epidemic regions, according to geographical closeness and infection level. In addition, a novel Lagrangian-SEIR-based model is posed for describing the dynamic within and between those regions. The capabilities of this methodology for identifying local outbreaks and reproducing the epidemic curve are discussed for the case of COVID-19 epidemic in Jalisco state (Mexico). The contagions from July 31, 2020 to March 31, 2021 are analyzed, with monthly adjustments, and the estimates obtained at the level of the epidemic regions present satisfactory results since Relative Root Mean Squared Error RRMSE is below 15% in most of regions, and at the level of the whole state outstanding with RRMSE below 5%.

**Keywords:** spatio-temporal dynamics; SARS-CoV-2; Lagrangian SEIR; regional heterogeneity; geographical network epidemic spread

---

### **1. Introduction**

The pandemic generated by the SARS-CoV-2 coronavirus (COVID-19) has caused major social, economic, and health impacts in the world, with more than 213 million confirmed cases; more than

4.95 million deaths [1] and a global economic growth fell 3.6 percent in 2020 [2].

In such a long and devastating pandemic, it has been necessary to adjust strategies and tools for better epidemiological surveillance. Typical strategies such as lockdowns, isolations, quarantines and border closures have predominated. They have often been applied in a similar and simultaneous way for a whole country or for a large territory, minimizing the fact that within same area there have been marked differences in the spread of the virus. This was the case of Mexico, where on March 30, 2020, the Federal Government ordered the suspension of non-essential activities throughout the country [3]. This is despite the fact that there are great contrasts in population concentration, ranging from the megalopolis of Mexico City to 184,000 small towns with less than 2500 inhabitants, many of them isolated from urban areas [4]. This has been one of the countries most affected by this pandemic where the leadership and management of the epidemic by the government has been highly questioned [5–7].

In these cases of large and diverse territories, a clear understanding of the spatio-temporal dynamics of COVID-19 is necessary to form the basis for decision making, planning and community action [8]. But it is important to emphasize the fact that this spatio-temporal perspective should seek to identify differences in regard how epidemic spreads across territory. This is because several studies have documented that the incidence of COVID-19 is related to the spatial and socioeconomic characteristics at the local level. For example, in [9] it is shown that geographical configuration, built environment densities, socioeconomic characteristics, and infrastructure are related to COVID-19 incidence in Germany when they are assessed at the county scale.

In the context of the COVID-19 pandemic, different methodologies still being proposed [10–13], and some of them argue the necessity of modelling the pandemic spread considering some type of geographical/spatial perspective. In [8] and [14] bibliographic reviews are made to summarize some of these proposals. In the case of [8] the main category considered in their review, was to locate studies of the temporal pattern of the COVID-19 phenomenon together with its geographic expansion. In [14], the relevant articles are classified and analyzed in three categories: disease mapping, exposure mapping, and spatial epidemiological modeling. This last category is the focus of this paper. This is in contrast to the fact that many published works only provide a representation of certain details in the form of a map, and in that sense they do not offer relevant contributions in the field of spatial analysis.

Approaches in literature based on models for describing the infection curve in a large territory or population, treating it as a whole isolated group are prevalent since are naturally formulated for capturing specific aspects of the epidemic with good results. See [15–17] for instance. However, their extension to capture heterogeneity on populations activity, or spatial dynamics could be not straightforwardly performed.

Although there are many papers that model spatial dynamics of pandemic spread, most of them develop from geographic units already given, whether these are countries, states, regions, counties or municipalities [18], and many of them describe the dynamic on the basis of a classical epidemic model (SEIR based models for instance) without particular considerations attending spread due to actual mobility or social interaction as driven terms on their models. See for instance works as [19–23]. In the case of [24] social aspects as age and race are considered, and corresponding regions are built in order to model the meta-population dynamics with SEIR-type model. However the interactions at region scale are just described as a procedure instead terms with epidemic significance. In [25] an age-stratified, nation-scale model is developed and social contact is modeled from Google Community Mobility data. Something that the COVID-19 pandemic has shown is that the dynamics of its expansion

has not respected the administrative divisions of the territory, but geographical borders of the expansion are closely related to the socioeconomic interconnection (that in place varies in time) instead.

In [26] a very complete review of the research activity in epidemic processes on networks is presented. The study of spatially structured interactions, such as between cities or more general geographical regions, is specially encouraged. Following the given advice, we model spatial component in a natural way by means of the construction of a suitable graph by means of Louvain algorithm. In [27] the authors presented quite interesting results about the classical final size relation of network epidemic models. In particular, they conclude that in complex networks with arbitrary degree distribution, the presented results result essential for estimating the expected magnitude or an outbreak severity of an epidemic. In our proposed method, the topology and size of the connection network remains optimal by construction. This graph allow our method to model spatial dynamics in a natural and realistic way in the context of the COVID-19 problematic.

In [28] the authors review recent efforts on modeling spatially explicit infectious diseases. Moreover, they show how modeling geographical regions, rather than a single, national approach, are quite likely to lead to better results. The authors in [29] show that infections, in particular the SARS-CoV-2, in one local district do not only depend on properties of the district, but also on the strength of social ties, mobility of population among districts and their own degree of infectiousness. It is clear that the virus does not stop at political borders and hence its spread through commute mobility must be studied. Furthermore, an updated review of GIS methodologies is presented in [30]. There, the authors make emphasis in that spatial analysis has been increasingly utilized to study the impacts of COVID-19. They present a very complete review of scientific results that used spatial science to study the pandemic in the second half of 2020. Their results point out that the use of spatial statistical tools have an upward trend.

There have been efforts in many directions to model propagation phenomena of a virus. Gandolfi in [31] presents the nontrivial interplay between epidemic and percolation models. The author proposed a new family of random networks which incorporate spatial features that are aimed to lead to a suitable spatial propagation modeling. As the own author remarks, such interplay between percolation and epidemics is still at an early stage and must be subject of further studies. In [32] the authors study the propagation of the ZEBOV virus behind West Africa's Ebola outbreak. There, they remark that the virus did not fundamentally change, but Africa did. Thus, they proposed stochastic models to study the geographic and socioeconomic effects that influence virus propagation. Although this study present similarities in the spatial mobility, the dynamics of infected people is remarkable different with the current pandemic of COVID-19. In [33] a complete study of incidence of AIDS and tuberculosis. In this paper, the authors show empirical findings that strongly contradict the assertions made by the National Research Council NRC, in 1993, saying that AIDS was going to remain confined within poor communities. The authors predict a fast spread over the New York metropolitan region showing that the density of commuting, as well as the intensity of poverty, completely determined the number of AIDS and TB cases per unit population. In all the above studies, the spatial interactions between regions shines as crucial in understanding any virus propagation phenomena.

Precisely this work focuses on posing a methodology in where spatio-temporal epidemic dynamics is captured by means of dividing a territory into time-varying *epidemic regions*. This division seeks to capture the mobility of the infectious agent by grouping neighboring geographical units (nodes) whose local dynamics of contagion spread are similar. Proper procedure for detecting these epidemic

regions is relevant for our study, then one of the contributions of this paper is an algorithm based on the Louvain algorithm [34] that couples infectious and geographical information together.

To describe the dynamics of the infectious disease, a system of differential equations that considers a *Lagrangian* movement [35] is proposed. This novel variation of a SEIR (Susceptible-Exposed-Infected-Recovered) type model, divides the population/territory into  $n$  epidemic regions and considers short temporal travel from one region to another. This formulation, in contrast to the classical SEIR models, considers open populations at region level and also the population distribution of the geographical area where the epidemic is taking place. Determining the value of  $n$  is a key aspect that is explored in the context of the methodology itself.

Thus, the objective of this paper is to detail the proposed methodology to divide the territory into epidemic regions, together with the Lagrangian model based on SEIR to describe the dynamics within and between those regions. The capabilities of this methodology for identifying local outbreaks, estimating exposed population and reproducing the epidemic curve are discussed for the case of COVID-19 epidemic in Jalisco state, which is one of the main states of Mexico.

This paper is organized as follows: Spatio-temporal pandemic spread study is motivated in Section 2.1. Our proposal model for regions, based on SEIR-type model with Lagrangian movement, is presented in Section 2.2. Methodology for identifying suitable regions for our model, is presented in Section 2.3. The study case for this work is addressed in Section 3.1. Finally, conclusions and discussion are presented in the light of our study case in Section 4.

## 2. Materials and methods

### 2.1. Epidemic regions

According to *Center for Disease Control and Prevention (CDC)*, infection of SARS-CoV-2, the virus that causes COVID-19, occurs by exposure to respiratory fluids carrying infectious virus. This exposure could occur in any of three ways: inhalation of very fine respiratory droplets and aerosol particles; deposition of respiratory droplets and particles on exposed mucous membranes (mouth, nose, or eye) by direct splashes and sprays (for instance, being coughed on); and touching mucous membranes with hands that have been soiled either directly by virus-containing secretions or indirectly by touching infected surfaces [36]. Thus, non-pharmaceutical intervention mechanisms (NPI), namely infected isolation, social distancing, socioeconomic activity reduction, travel restrictions, among others, are intended to decelerate the infection curve at human-human contact scale. Recent contributions related to effectiveness analysis of NPI's on different world regions can be found on [37–41].

It is known that disease spread between cities is driven by population mobility and also, within a specific city, disease dynamics is affected by local factors. Thus spatial heterogeneity is commonly considered for studies at region, state or municipality scale (see [42–44]) and those factors are directly related to socioeconomic activity of inhabitants. In other words, two connected cities or population groups, with highly close socioeconomic activity, will present very similar outbreak dynamics and could be considered as a single group. This rationale may be seen as equivalent to the significant factors observed on [24].

Knowledge concerning spatio-temporal evolution of epidemics, within and between regions, helps in assessment of currently available/depleted resources, and forecasting of needed resources or

actions, derived from public health policies. Classical compartment-based models (see [35, 45]), computational models based on networks or graphs representing groups (see [11, 18–20, 22] for recent works modelling mobility and spatial spread, and [23, 24, 39, 44, 46, 47] for network models), are naturally employed for this aim. But as mentioned above, it is important to consider whether nearby populations share similar contagion dynamics, so that at a given time they form part of the same epidemic region. Again these regions will not necessarily coincide with administrative or economic regions.

Capabilities of models based on epidemic dynamics per region as diagnostic instrument, resort at first instance on the correct regions definition. As second instance, on a suitable model that assembles those regions, and that might be properly implemented/codified into a numerical procedure that describe the epidemic dynamics. This problematic, to our knowledge has not been integrally attended on literature. The methodology proposed address both aspects: first, epidemic regions are identified according infection level and geographical connection together; and second, multiple SEIR-type models are placed to describe the inner dynamics of geographical units and contagions by mobility. These aspects are described below.

## 2.2. Epidemiological model for regions

First option for epidemiological mathematical modelling is the classical *Susceptible-Infected-Recovered (SIR)* compartmental model posed by [48]. However, it has been remarked on literature concerning COVID-19, the relevance of asymptomatic or exposed population onto dynamic of the epidemic. Thus, it is recommended *Susceptible-Exposed-Infected-Recovered (SEIR)* compartmental model, defined by the ordinary differential equations system:

$$\frac{dS}{dt}(t) = -\beta \frac{S(t)I(t)}{N}, \quad (2.1)$$

$$\frac{dE}{dt}(t) = \beta \frac{S(t)I(t)}{N} - \sigma E(t), \quad (2.2)$$

$$\frac{dI}{dt}(t) = \sigma E(t) - \gamma I(t), \quad (2.3)$$

$$\frac{dR}{dt}(t) = \gamma I(t), \quad (2.4)$$

with initial conditions

$$S(t_0) = S_0, \quad E(t_0) = E_0, \quad I(t_0) = I_0, \quad R(t_0) = R_0, \quad t_0 \leq t \leq t_0 + T, \quad (2.5)$$

where total population  $N = S(t) + E(t) + I(t) + R(t)$  remains constant, parameter  $\beta > 0$  corresponds to transmission rate,  $\sigma > 0$  to incubation rate, and  $\gamma > 0$  to recovery rate and  $T$  is an user parameter that corresponds to the final time chosen for the study.

It has been noticed that pandemic spread across a large territory is composed by multiple local outbreaks, appearing on different times, and spreading at velocities varying from region to region (highly heterogeneous). Naturally, spread within region adopts its particular dynamic that may be described by SEIR model (Eqs (2.1)–(2.5)). Thus, two mobility phenomena are in consideration. The first one is present within regions, the same that supports the epidemic outbreak and is captured by SEIR model per region. The second one, that is present between geographically connected (or

neighbours) regions. The later is causative of epidemic spread across larger areas such as states or countries and can be seen as immigration term in open SEIR model for multiple regions.

Short term mobility between regions is described by *Lagrangian* movement model. It is assumed total study population  $N$  is divided into  $n$  regions, such as  $N = N_1 + \dots + N_n$ . For each region  $i \in 1, \dots, n$ , it is satisfied  $N_i = S_i + E_i + I_i + R_i$ . Susceptible individuals on region  $i$  are getting infected when visiting any region  $j \neq i, j = 1, \dots, n$ , by infectious individuals residing on region  $j$ . Then, transmission term on SEIR model (Eqs (2.1) and (2.2)), from Lagrangian mobility consideration described on [35], becomes the sum:

$$\beta \frac{S(t)I(t)}{N} \longrightarrow S_i(t) \sum_{j=1}^n \beta_{ij} \frac{I_j(t)}{N_j}. \quad (2.6)$$

Here it is assumed that all regions are mutually connected. Transmission within a region corresponds to  $j = i$  on Eq (2.6), and transmission by mobility is given by  $j \neq i$ . In this formulation, transmission rates try to explain contagions due to mobility. For instance,  $\beta_{kl}$  should be zero on unconnected regions even if both regions  $k, l$  are highly infectious. Otherwise,  $\beta_{kl} \neq 0$  would induce artificial infections assuming not existent mobility. This double role of  $\beta_{kl}$  parameters (namely describing connection, and transmission rate within and between regions), often conducts to misleading interpretations of transmission rate. For example, it may not be distinguishable whether low values of  $\beta_{kl}$  obey to low mobility rates (connectivity in turn) or low infection causality between neighbour regions. Even more, this parameters may be significantly different from the inner transmission rate for each region.

Concerning the later point, we propose to maintain an overall transmission rate  $\beta$ . The rationale is to separate transmission rate  $\beta$ , as the underlying characteristic of epidemic that is independent of any region, whereas connectivity and infection causality due to mobility is computed from existing connection between regions and actual infection level. This is explained below.

Let us remark that infection process due to inner mobility for a region differs from outer mobility for same region. With this in mind, let us split sum on Eq (2.6) for analyzing inner dynamic of regions  $i$  separately from mobility terms. Then, we have,

$$S_i(t) \sum_{j=1}^n \beta_{ij} \frac{I_j(t)}{N_j} = S_i \beta_{ii} \frac{I_i(t)}{N_i} + S_i(t) \sum_{j=1, j \neq i}^n \beta_{ij} \frac{I_j(t)}{N_j}. \quad (2.7)$$

Later, from these considerations above, we propose contact rates given by:

$$\beta_{ij} = \begin{cases} \alpha_i \mathbb{I}_{i,j} \frac{I_i(t) + I_j(t)}{S_i(t)}, & \text{for } i \neq j, \\ \beta, & \text{for } i = j. \end{cases} \quad (2.8)$$

where  $\alpha_i$  is a mobility rate. The rationale of Eq (2.8) is the following: infection due to mobility requires that a susceptible individual, chosen with probability  $1/S_i(t)$  at time  $t$  in region  $i$ , encounters an infected individual. This infected individual may come from a neighbouring region  $j$  (with rate  $\alpha_i I_j(t)$ ), or from same region (with rate  $(\alpha_i I_i(t))$ ). Term  $\mathbb{I}_{i,j}$  is an indicator function that is equal to 1 when there exists a geographical connection between regions  $i$  and  $j$ , and 0 otherwise. As result, expression for  $\beta_{ij}$  may be computed at any time and codifies existing communication between regions (in  $\alpha_i$  and  $\mathbb{I}_{i,j}$  terms) and the actual infection level on both communities.

Equation (2.8) provides information concerning to diffusion of infection between regions in a more natural way than Lagrangian terms of Eq (2.6). This is, in the case of a small community  $i$ ,  $I_i(t)$  and  $S_i(t)$  would be also small, but when  $i$  is neighbour of a region with high number of infections, the chances that a infected individual arrives to region  $i$  are also high. Moreover, since  $S_i(t)$  appears dividing on Eq (2.8), the effect of infection due to immigration will have more impact on region  $i$ . When region  $i$  is not a neighbour of a highly infected region, infection diffusion does not occur due to indicator function  $\mathbb{I}_{i,j}$  will be equal to zero.

Replacing Eq (2.8) on Eq (2.7) gives us:

$$\begin{aligned} S_i \beta_{ii} \frac{I_i(t)}{N_i} + S_i(t) \sum_{j=1, j \neq i}^n \beta_{ij} \frac{I_j(t)}{N_j} &= \\ &= S_i \beta \frac{I_i(t)}{N_i} + S_i(t) \sum_{j=1, j \neq i}^n \alpha_i \mathbb{I}_{i,j} \frac{I_i(t) + I_j(t)}{S_i(t)} \frac{I_j(t)}{N_j}, \\ &= S_i \beta \frac{I_i(t)}{N_i} + S_i(t) \frac{\alpha_i}{S_i(t)} \sum_{j=1, j \neq i}^n \mathbb{I}_{i,j} (I_i(t) + I_j(t)) \frac{I_j(t)}{N_j}, \\ &= S_i \beta \frac{I_i(t)}{N_i} + \alpha_i \sum_{j=1, j \neq i}^n \mathbb{I}_{i,j} (I_i(t) + I_j(t)) \frac{I_j(t)}{N_j}. \end{aligned}$$

Since term  $\mathbb{I}_{i,j}$  indicates geographical connections between regions, it is defined a connectivity matrix  $W^t = [w_{ij}^t]_{i,j=1}^n$ , where entries are defined by:

$$w_{ij}^t = I_i(t) + I_j(t), \quad (2.9)$$

and diagonal terms  $w_{ii}^t = 0$ . For practical purposes, in our implementation, an additional simplification is made on weights  $w_{ij}^t$ . It will be assumed that infection level for  $I_i$  and  $I_j$  will not change drastically in short intervals  $t \in T^k = [t_k, t_{k+1})$ . Then, weights are computed at the beginning (on time  $t_k$ ) and remain constants on  $T^k$ . Then, weights should be updated again on time  $t_{k+1}$ . This simplification is not a drawback, since weights can be updated once  $I_i$  and  $I_j$  are computed. Note that in consequence, the regions structure is assumed constant within time interval  $T^k$ . This will be described later.

Then, our corresponding SEIR model with Lagrangian movement, for each region  $i \in 1, \dots, n$  is given by:

$$\frac{dS_i}{dt}(t) = -\beta \frac{S_i(t)I_i(t)}{N_i} - \alpha_i \sum_{j=1, j \neq i}^n w_{ij} \frac{I_j(t)}{N_j}, \quad (2.10)$$

$$\frac{dE_i}{dt}(t) = \beta \frac{S_i(t)I_i(t)}{N_i} - \sigma E_i(t) + \alpha_i \sum_{j=1, j \neq i}^n w_{ij} \frac{I_j(t)}{N_j}, \quad (2.11)$$

$$\frac{dI_i}{dt}(t) = \sigma E_i(t) - \gamma I_i(t), \quad (2.12)$$

$$\frac{dR_i}{dt}(t) = \gamma I_i(t), \quad (2.13)$$

with initial conditions

$$S_i(t_0) = S_{i,0}, \quad E_i(t_0) = E_{i,0}, \quad I_i(t_0) = I_{i,0}, \quad R_i(t_0) = R_{i,0}, \quad t_0 \leq t \leq t_0 + T. \quad (2.14)$$

Here, terms  $\beta$ ,  $\gamma$  and  $\sigma$  are the corresponding transmission, recovery and incubation rates (as in SEIR model),  $\alpha_i$  is mobility rate and  $w_{ij}$  are infection connectivity weights defined in Eq (2.9). We refer to Eqs (2.10)–(2.14) as *Epidemic Regions SEIR (ERSEIR)* model.

Let us remark the relevance of infection connectivity weights  $w_{ij}$  in contrast with original Lagrangian mobility Eq (2.6). In Eq (2.6), transmission rates  $\beta_{ij}$  becomes variables that should be estimated from data. This leads to large dimensional problems where suitable numerical implementations and prior information are mandatory. In our case, weights  $w_{ij}$  are computed from infection levels of current solution of ODE solver, together with epidemic regions identification procedure described below. Consequently, the total number of free parameters in our formulation is  $n + 3$ , for  $n$  regions. Namely,  $n$  parameters are for mobility rates  $\alpha_{i=1}^n$ , and 3 more corresponds to  $\beta$ ,  $\gamma$  and  $\sigma$ . In other words, a single value of  $\beta$ ,  $\gamma$  and  $\sigma$  is used for all epidemic regions, and spread between epidemic regions is conducted by weights  $w_{ij}$ . In addition, all these parameters have epidemiological interpretation.

There have been some interesting studies of the previous dynamics. For instance, in [49] an asymptotic analysis of multi-patch stochastic SIS model is studied. They present remarkable conclusions about the behaviour of the model when the number of sites and individuals goes to infinity. However, it is not straightforward to add further model assumptions like the ones related to social and economical aspects on the COVID-19 propagation. In [50] the authors derive a multi-patch model describing the propagation of a virus within a structured host population. It is assumed that the virus propagates by direct transmission and by indirect contamination of susceptible through the environment. Due to the nature of the virus, once infected, individuals do not recover from the disease. This last assumption is not share with the nature of the virus we are studying in this work. In [51] a spatial propagation model of a disease between multiple species is proposed. There, the spatial component is modeled by the use of patches. The authors show simulations for the spread of a disease in one species and only two patches. In contrast, we deal with many more spatial regions represented by a suitable constructed graph. In [52] the dynamics of a multi-patch SIR model are used to study disease persistence and extinction problems. In [53] a general SEIRS multi-patch and multi-group model is proposed. There, they consider differential state-host mobility patterns, by a Lagrangian approach, to investigate the effects of heterogeneity in mobility patterns. In most of related multi-patch SEIR epidemic models, epidemic dynamics is studied by exploration of previously established patches and/or population units (groups defined by ethnicity or activity for instance). In this work, population units, specifically the epidemic regions, and their transmission rates are build within methodology itself as result of infection level and geographical information together.

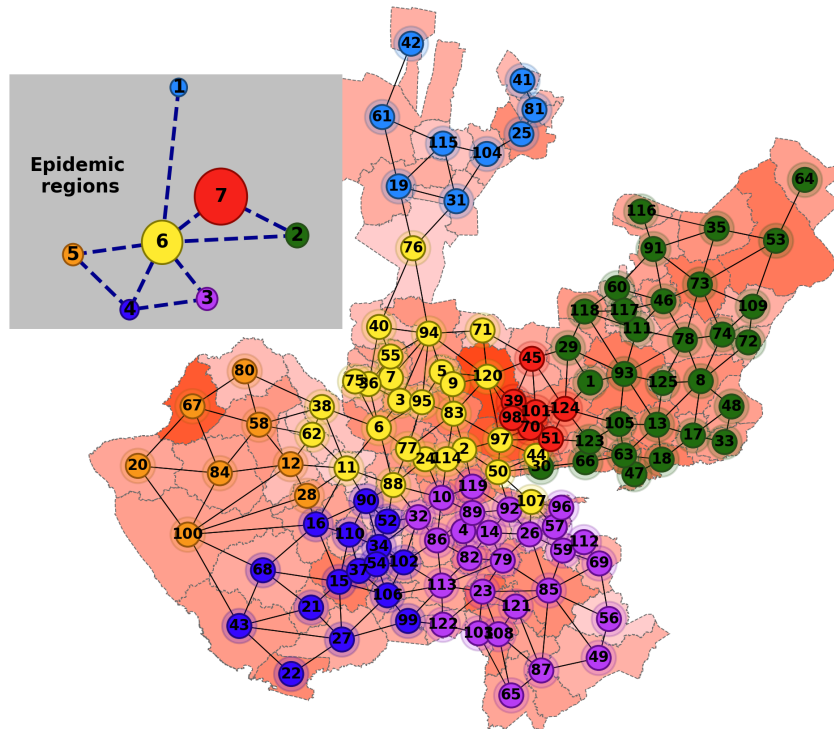
### 2.3. Methodology for identification of epidemic regions

Let us consider an undirected graph representing a geographic region. See graphs on Figure 1 for instance. Nodes on graph represent geographic units, as states, cities or municipalities, and edges between nodes indicate the existence of direct connection between two geographical units by roads. This is, edges indicates neighborhood.

Epidemic regions target is exemplated around orange nodes on Figure 1. Geographical units related to orange nodes present similar infection level and are geographically connected. In addition, node 67 is leading the outbreak in that region. Nodes 17, 68 and 43 are colored in blue due to their infection seems to be more influenced by a local outbreak raised on node 15. This behavior is some similar in the



interface between other regions. Systematical epidemic regions identification with those considerations can be seen as a clustering or community detection problem.



**Figure 1.** Graph representing the 125 municipalities of Jalisco state (the case study). Edges indicate direct connection between two neighbour municipalities by roads, and corresponding weight of connection is computed from the infection level between both municipalities. Numbers are used to index the municipalities and colors to identify epidemic regions. Epidemic regions are represented by a 7-regions graph on the upper left corner of Figure. Color intensity of each city is proportional to number of confirmed active cases at July 31, 2020.

Some algorithms for communities detection are based on a hierarchical approach, where nodes are joined from a priority function. In [54] it is presented the Girvan-Newman algorithm where the concept of *betweenness* is introduced. From that seminal contribution, further methods have been developed, for instance Label propagation [55], Fluid communities [34] and Louvain algorithm. The later is used in this work for reasons explained below.

Target of community algorithms is to build a graph partition  $G_n$  with  $n$  communities, that divides a simple graph  $G(V, E)$ , with  $V$  nodes and  $E$  edges, maximizing the *modularity* function:

$$Q(G_n) = \frac{1}{2L} \sum_{e_{rs} \in E} \left[ \hat{w}_{rs} - \frac{k_r k_s}{2L} \right] \delta(c_r, c_s), \quad (2.15)$$

where  $L$  is the sum of weights of edges on the graph,  $[\hat{w}_{rs}]$  is weighted adjacency matrix:

$$[\hat{w}_{rs}] = \begin{cases} \hat{w}_{rs}, & \text{weight for edge connecting node } r \text{ and node } s, \\ 0, & \text{no connection,} \end{cases}$$

$e_{rs}$  indicates the edge connecting nodes  $r$  and  $s$ ,  $k_l = \sum_s \hat{w}_{ls}$  is the sum of weights of incident edges to the node  $l$ ,  $c_r$  indicates current community for node  $r$  and  $\delta$  is the delta function:

$$\delta(c_r, c_s) = \begin{cases} 1 & \text{if node } r \text{ and node } s \text{ are in same community,} \\ 0 & \text{otherwise.} \end{cases}$$

The term  $\frac{k_r k_s}{2L}$  is considered as a probability of existence of an edge between nodes  $r$  and  $s$  according their respective weight degree.

Let us point out that weights for spread between epidemic regions are taken as the weights  $w_{ij}$  indicated on Eqs (2.10)–(2.13), and they should not be confused with the weights between nodes (geographical units) denoted by  $\hat{w}_{rs}$ . The interpretation on both cases is similar, however the context of applicability changes. Within Louvain algorithm, community nodes are collapsed into a single representing community node. Corresponding inner weighted edges are added in a single self-connecting weighted edge. Equivalently, connections between nodes in different communities are added as a single inter-community weighted edge. Those actually could be used as the weights  $w_{ij}^t$ . However, the interpretation of weights in that context turns not so natural. In place, we opt to take the weights based on weight of node (namely infection level) instead edge, in order to maintain an interpretation as infection level. Corresponding weights are described below. Dynamics within each epidemic region may be modeled with the help of a standard SEIR model with mobility.

Rationale of modularity function is to promote that communities are defined by clusters of nodes densely intra-community connected whereas number of inter-community connections is lower than number of connections on a equivalent randomly connected graph. Negative or zero values of modularity, means that a community structure cannot be identified from given graph (a fully connected graph for instance). Positive values of modularity indicates that a community structure is identified and upper bound for modularity is 1. In this sense, the large modularity value is, the better quality of the community structure. In [56] it is mentioned that optimal partition may be indistinguishable from other partitions with high modularity value. As result, optimal partition should be identified by addition of particular information. Thus, modularity is application dependant, and in practical applications, a good modularity is obtained from experience. In [57] it is computed a modularity value for a graph partition obtained onto a large scale hierarchical graph. They report in that context  $Q(G_p) = 0.3$  as good modularity value. In this work, we employed the Louvain algorithm implementation of Gephi software [58].

In contrast with natural approaches where weight of connection is defined by connectivity level, in our model weights are computed as in Eq (2.9) at geographic units instead epidemic regions. That is the weight of one edge connecting node  $r$  and node  $s$ , at time  $t$  is computed as:

$$\hat{w}_{rs}^t = I_r(t) + I_s(t). \quad (2.16)$$

This approach produces a weighted undirected graph, that gathers geographical connection and infection level at once. Thus, epidemic spread across geographic units occurs from highly infected units

towards surrounding units only if there exist mobility between them. Note that, since weights from Eq (2.16) vary with time, the structure of communities will change with time according infections levels.

It is important to mention that, since Louvain algorithm is designed to increase modularity by selecting nodes for each community in such a way that density of graph is scattered (together with infection levels on Eq (2.16)), it turns out that highly infected neighbouring geographic units (nodes with high weights) are promoted to be in different communities. That is, once a unit increases its infection level significantly due to a local outbreak, a new community is created to restrain the density level of the graph, and its epidemic dynamics is therefore captured by this resulting graph.

In contrast with Louvain algorithm, by means of using the betweenness approach of the Girvan Newman algorithm, it could be identified the geographic units that play important role (as a bridge between regions) on the epidemic spread across state. This will be part of future studies.

With the consideration that epidemic dynamics of a node belonging to specific community is represented by the epidemic dynamics of the entire community, the dynamics of the entire state is described by our ERSEIR model with the number of epidemic regions equals to number of communities  $n$ , as first hierarchical diagnostic level.

Regarding the hierarchical structure in which the spread of an epidemic can be analyzed, some remarks may be indicated. It is a fact that hierarchical diffusion models are relevant for studying epidemic spread in modern times [59]. Since COVID-19 expansion has not respected administrative divisions of territory, neither at municipalities nor country scales, our methodology as is applied in this work is limited to a scale of municipalities within single state, and could be applied on a more general scale. However, the methodology is not covering interaction at neither inter-country nor intra-country scale due to is not the main aim of our study. However, such an interactions could be incorporated in one of two simplified ways. First one, suggested by anonymous reviewer, comes from a block-modelling approach, in which blocks of non-zero (low valued) weights are placed off of the diagonal of the (reordered) adjacency matrix, for incorporating “peripheral” relations (neighboring states or countries, or connections with other non-neighboring geographic entities) whereas relations obtained from our epidemic regions procedure are maintained as blocks along the diagonal of the later adjacency matrix. In other words, outer-interactions of state are represented as a block of nodes communicating inner nodes. In this scenario, we envisage that in cases of peripheral blocks densely intra-connected, Louvain algorithm will build such a block as a sole community naturally, and then ERSEIR methodology is straightforwardly applicable. Second one, consist in considering a (surely very large) initial graph formed by every geographical unit within a country for instance. Consider the graph shown on Figure 1, for such a initial graph, Louvain algorithm will build an equivalent graph (as the one shown in small panel of Figure 1) composed by the epidemic regions within country. Then, infection by interactions at next hierarchical level (countries for instance) may be added as epidemic region nodes. Next step, could estimate the epidemic dynamics from a modified super version of Eqs (2.10)–(2.14) model, or even consider a next hierarchical level formed by groups of epidemic regions that in turn are groped by Louvain algorithm, and so on towards upper hierarchies. In this second approach, Eqs (2.10)–(2.14) should be modified as an open version, to consider the inter-communities contagions and be able to propagate the external infectious across hierarchical levels. In both cases: the block-modelling approach and super ERSEIR, an interesting computational problem arises due to increase on the problem dimension. Relevant aspects of numerical implementation of ERSEIR model are lead to be described in the context of the case of study in next section.

### 3. Experiments and results

#### 3.1. Case study: Jalisco

Diagnosis capabilities of ERSEIR model are analyzed in the case of Jalisco state within the period ranging since July 31, 2020 to March 31, 2021. The interest of this case of study lies in the fact that Mexico is one of the most affected countries, with more than 286.259 confirmed deaths as of 10/24/21 [60], and a fall of 8.2 percent in its gross domestic product in 2020 [2]. Jalisco is one of the main states of Mexico, with 8.4 million inhabitants, distributed in 125 municipalities. Many of them are small rural municipalities, but also include medium-sized cities, such as Puerto Vallarta, and a large city: the metropolitan area of Guadalajara. The later has 5.3 million inhabitants, a good connectivity with the rest of the country, and with the United States and Canada. In this way, due to these characteristics of Jalisco state, it is interesting to characterize its the epidemic dynamics. In addition, the 125 municipalities of Jalisco are officially grouped into 12 administrative regions, which were established according to their closeness and similarities in terms of factors such as socioeconomic characteristics, hydrological basins, road connectivity, among others [61]. So it is interesting to contrast the epidemic regions with these administrative regions. We refer to the later as economic regions.

This case of study is depicted as the graph with edges in solid black lines on Figure 1. In this graph the 125 municipalities of Jalisco state are represented as nodes, where edges between nodes indicate the existence of a common land border between two municipalities. Graph on blue dashed edges corresponds to epidemic regions, resulting from procedure described in Section 2.3 with the following considerations.

Weights of connection (Eq (2.16)) for graph of Figure 1, are computed from the confirmed active cases on July 31, 2020 for each node (color level of municipalities on Figure 1 is proportional to this value). We take then initial time  $t_0$  at July 31, 2020. Thus, each entry of the matrix of initial weights is defined by:

$$\hat{w}_{rs}^0 = I_r(0) + I_s(0).$$

With this weights, the Gephi implementation of Louvain algorithm constructs  $n = 7$  communities (indicated by colors on Figure 1) with a resulting modularity of  $Q = 0.389$ . At the same time, Louvain's algorithm provides the corresponding weighted connectivity matrix  $w_{ij}$ . That is, Jalisco state is divided on  $n = 7$  epidemic regions at the beginning of this study.

Within each epidemic region  $i = 1, \dots, n$ , susceptible, exposed, infected and removed initial populations are estimated as follows.

- Susceptible:

$$S_{i,0} = N_i - I_{i,0} - E_{i,0} - R_{i,0},$$

where  $N_i$  is computed as the sum of inhabitants in cities belonging to corresponding epidemic region  $i$  (source [4]) and is considered fix along time.

- Exposed:

$$E_{i,0} = C(t_0 + 6) - C(t_0),$$

where  $C(t_0)$  corresponds to number of active cases\* reported by [60] at time  $t_0$ , and  $C(t_0 + 6)$

\*Data base considers date from first symptom, rather that confirmed active case date. However this specification is transparent for the methodology.

indicates the active cases 6 days after initial time. This consideration is consistent with the *age time*, or incubation period of virus after exposition (6 days on average according WHO and [62]). Note that this value is required only for initial condition estimates, and therefore does not restrict the application our methodology for  $t > t_0$ .

- Infectious:

$$I_{i,0} = C(t_0) - C(t_0 - 14),$$

where  $C(t_0 - 14)$  is the number of active cases 14 days before  $t_0$ . Since 14 days is the time for recovering from COVID-19 infection, the rationale is to consider that some cases are getting recovered at time  $t_0$ .

- Removed:

$$R_{i,0} = C(t_0 - 14).$$

This removed category corresponds to the cases subtracted from infectious population.

Parameter values of ERSEIR model, are placed according to WHO reports. This is, recovery rate was left as  $\gamma = \frac{1}{14}$  according to 14 days of recovery mentioned above, and it keep fixed since is an inherent characteristic of the pathogen. For estimating the value of  $\beta$  (Eqs (2.10) and (2.11)), we consider that according to studies [63–65],  $\mathcal{R}_0$  is close to 1.4, but it can even go up to 4.2. Then given  $\gamma, \beta$  could be in range  $\beta \in (0.01, 0.3)$ . Parameter  $\beta, \alpha_i$  and  $\sigma$  were manually adjusted every period to fit the data<sup>†</sup>. Thus, parameters for ERSEIR were taken as:

$$\beta \in [0.01, 0.10713], \quad \gamma = 0.07142, \quad \sigma \in [0.1, 0.25]$$

$$\alpha \in [0.01, 1].$$

In the case of parameter  $\sigma$ , the range for manual fit was taken from [66]. A time step of one day was considered in our experiments.

On Figure 2 are presented the curves (in red) corresponding to estimated infectious cases for each one of the 7 epidemic regions, obtained from our ERSEIR model implementation and for Jalisco state. Actual confirmed cases obtained from [60] data are drawn on blue lines with circular marks. These data correspond to sum of confirmed cases in the municipalities belonging to every epidemic region. It is observed that relative error (drawn on green line with squared marks) is below of 20% in most of the estimations. In addition, in the cases with high values of relative error, the actual difference in number of cases estimated against data (black numbers over black dashed lines) is low. Since each epidemic region has different scale, quantitative evaluation of the per-epidemic region quality of the estimation should not be performed via *Mean Squared Error (MSE)* (see [67] for further details). In place, a scaled-independent monthly<sup>‡</sup> comparison is performed by the *Relative Root Mean Squared Error (RRMSE)* computed as:

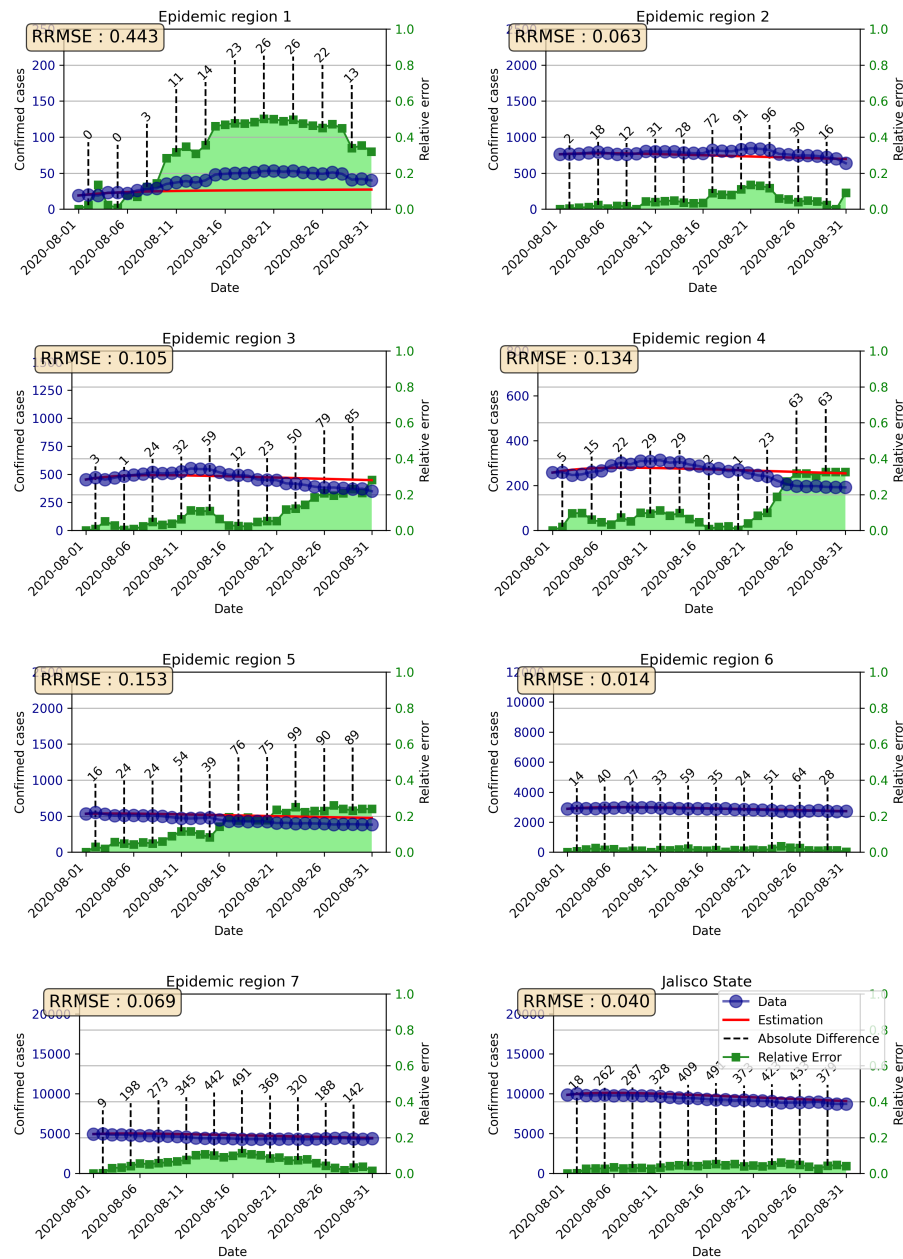
$$RRMSE_i^m = \frac{\sqrt{\frac{1}{M} \sum_{t=1}^M (I_{i,t} - d_{i,t})^2}}{\frac{1}{M} \sum_{t=1}^M d_{i,t}}, \quad \text{for } i = 1, \dots, n,$$

where  $m$  corresponds to month,  $M$  is for averaging over number of days in month  $m$ ,  $I_{i,t}$  are the infectious for epidemic region at day  $t$ , and  $d_{i,t}$  is the corresponding COVID-19 infectious data. It is

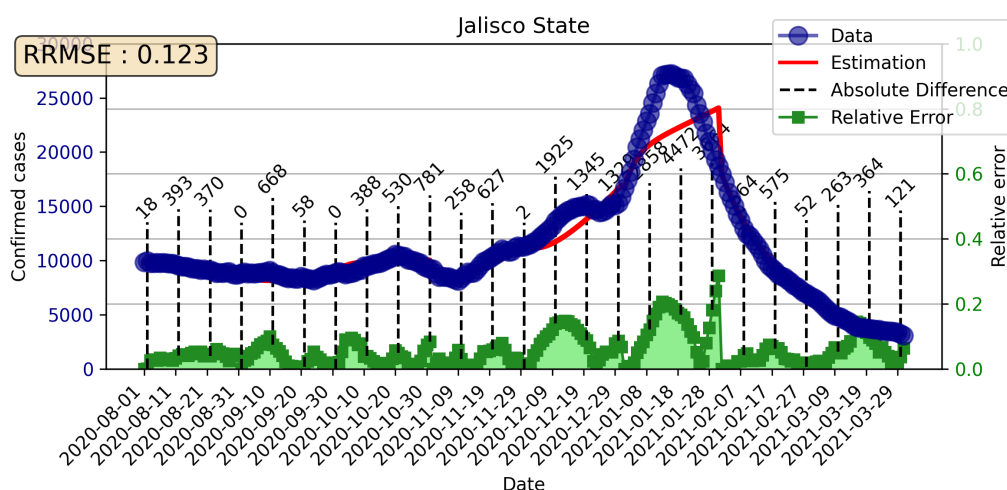
<sup>†</sup>Applying methodologies for estimating parameters from data with this methodology will be considered in further studies

<sup>‡</sup>It is chosen monthly computations of RRMSE since it is the time up to epidemic regions are recomputed in our experiments.

observed that RRMSE values for August (upper left box on every panel on Figure 2) is consistently lower than 20% with respect data, with exception of epidemic region 1, where the magnitude of data is very low. It is remarked that better estimations are obtained on larger regions and total Jalisco state cases; with an excellent RRMSE.



**Figure 2.** Daily estimating results from August 1, until August 31, 2020 obtained from ERSEIR implementation are presented in red lines, for each epidemic region and for Jalisco state. Active cases data (source [60]) are drawn on blue line with circular marks. Relative error is computed for every day and drawn on green with squared marks and values corresponds to right axis. Numbers over dashed black lines indicates the absolute error.



**Figure 3.** Estimates obtained from ERSEIR model from August 1, 2020, until March 31, 2021, are drawn on red line. Active cases data (source [60]) are drawn on blue line with circular marks. Relative error is computed for every day and drawn on green with squared marks and values corresponds to right axis. Numbers over dashed black lines indicates the absolute error.

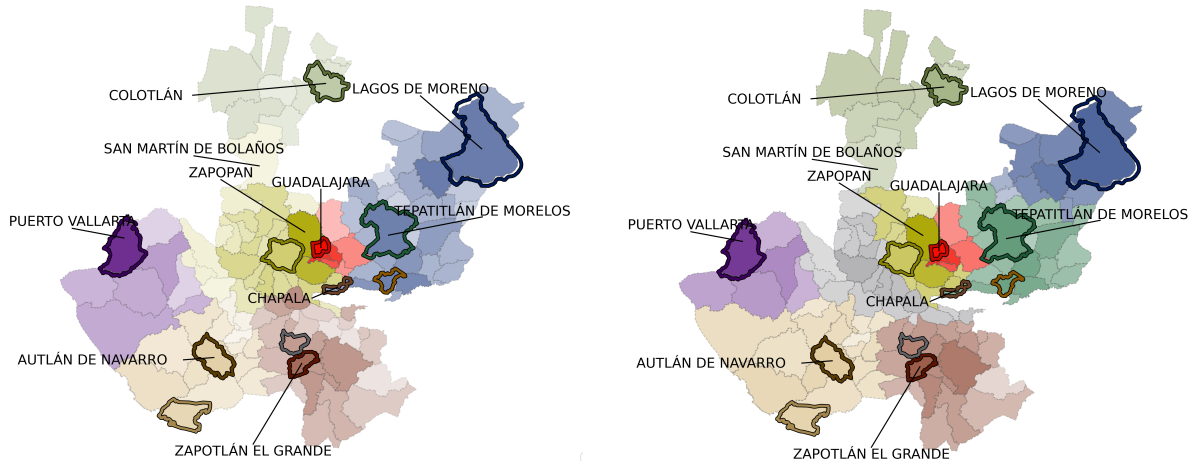
On Figure 3 are shown the estimates from ERSEIR model for time ranging from August 1, 2020, until March 31, 2021. These estimates are obtained as the sum of infectious of every epidemic region. This is:

$$RRMSE = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (I_{\cdot,t} - d_{\cdot,t})^2}}{\frac{1}{N} \sum_{t=1}^N d_{\cdot,t}},$$

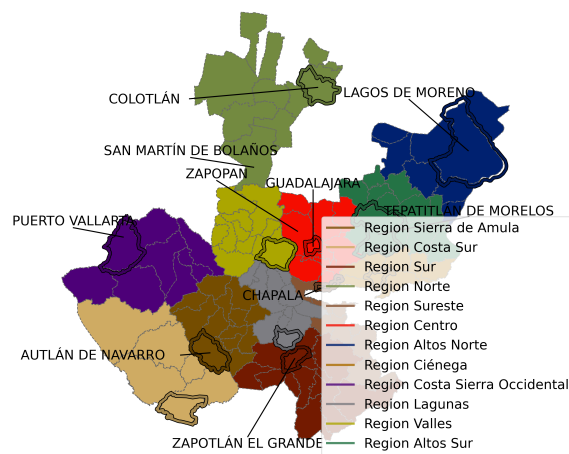
where  $\cdot$  indicate that is not dependent of epidemic region. It is observed that the relative error is below 10% most of the time and the RRMSE is also reasonably low: 0.123. In addition, it is interesting to note that some jumps appear at the end of months October, December and January. Around that months, the pandemic peaks and in addition, the increase of confirmed cases obeys to socioeconomic activity increase due to end-of-year festivities. Moreover, NPI and public health policies are not being considered in ERSEIR model. We envisage that better estimates may be obtained by adding NPI considerations to ERSEIR model, and reducing time between epidemic regions identification step (perhaps after 15 days of simulation instead a month). Later aspects are in considerations for future work.

On Figure 4 are presented the epidemic regions identified from Louvain algorithm implementation. Panel (a) shows the seven initial epidemic regions, obtained when considering initial numbers (on July 31, 2020) and it is the same division that is presented on Figure 1. With the numbers of infectious at the end of August 2020, weights on Eq (2.16) are recomputed and Louvain algorithm is executed again to reproduce the epidemic dynamics of September, 2020. In consequence, epidemic regions are redefined for next month. This procedure is repeated every month until January 31, 2021 and it produces the nine epidemic region division shown on panel (b). January month is chosen due to corresponds to month when pandemic peaks and coincidence between epidemic and economic regions is more evident. Color of municipalities on panels (a) and (b) indicate belonging epidemic

region and intensity level of color is proportional to number of infectious from the ERSEIR implementation with respect to total active cases at the end of the month. On panel (c) it is presented the current division of Jalisco state based on economic activity and geographic criteria [61]. Colors of panels (a) and (b) where chosen as colors of epidemic regions to associate the similarities in both divisions. The municipalities with highest economic activity on each region of panel (c) are highlighted. Same leading municipalities per region are also highlighted on panels (a) and (b) for identifying in which epidemic region are located.



(a) Seven epidemic regions on July 31, 2020. (b) Nine epidemic regions on January 31, 2021.



(c) Twelve economic regions

**Figure 4.** Division of Jalisco state on epidemic regions according to Louvain algorithm and active cases at July 31 (panel (a)) and August 25, 2020. Current division in regions for state, according to economic activity and geographic criteria, is shown on panel (c).



There are some remarks to do around Figure 4. The first one is the relation between social-economic activity and epidemics dynamics is evidenced. It is observed that municipalities: Tepetitlan de Morelos, Colotlán, Guadalajara, Puerto Vallarta and Zapotlán el Grande, have the largest number of confirmed cases on epidemic regions of panel (a). Those are also the leading municipalities on the economic regions of panel (c). In the case of Lagos de Moreno on July 31, it is assigned to epidemic region of Tepetitlan de Morelos, see panel (a). The rationale is, socio-economic activity between Tepetitlán de Morelos and Guadalajara is more active, due to closeness and the fact that Guadalajara is the capital city. Thus epidemic spread is delayed for Lagos de Moreno. However, once infection evolves, Lagos de Moreno develops its own local outbreak that is captured by Louvain algorithm, dividing blue region on panel (a) into two epidemic regions (blue and green on panel (b)). Resulting epidemic regions are consistent with economic regions of panel (c) (on green and blue respectively).

As natural, since Guadalajara city is the capital of state, it has largest numbers of inhabitants and economic activity. Thus, it governs the outbreak in municipalities belonging Region Centro. However, it is remarkable that its neighbour municipalities of Zapopan is grouped (and kept) on its own epidemic region. This is explained from the fact that Zapopan has enough inhabitants and inner economic activity for supporting its own outbreak and in turn, influences surrounding municipalities. In fact, on July 31, the municipalities of San Martín de Bolaños appears to be more influenced by the local outbreak of Zapopan (see panel (a)) and this is natural since San Martín de Bolaños is a bridge city between Region Norte and the rest of the state. However, on January 31, 2021 (panel (b)) its epidemic region changes due to increment of infectious in the northern cities (specifically in Colotlán) and the existent economic activity, as result epidemic region and economic region coincide.

Another interesting case is the municipalities of Chapala. It is the leading city of Region Sureste (see panel (c)). However, there exists a lake between Chapala and the rest of the municipalities belonging this economic region. This conform a natural barrier that diminishes the social activity. Then, Chapala is grouped in epidemic regions where there exists a more direct social activity due to road connections.

In the case of Puerto Vallarta, it expected that it govern the outbreak in surrounding municipalities, since Puerto Vallarta concentrates most of the socio-economical activity in its region due to tourism.

Regions Laguna, Costa Sur, and Sierra de Amula, are conformed by towns and settlements with a considerable degree of marginalization. Thus, it is reasonable that there is not a clear leading city in terms of infectious and therefore the epidemic regions are wider than corresponding economic regions. In other words, since there is not a clear leading municipalities, or local outbreak (see that intensity level of color around that region is almost uniform) the corresponding epidemic region is extended (see both: panels (a) and (b)).

Due to space constraints, temporal evolution of epidemic regions every month is presented as supplementary material.

Finally, it should be noted that our model is able to provide an estimation for exposed population. However, this was not shown in this work due to impossibility for having data for verification.

#### 4. Conclusions

This work addresses the problematic of describing the spatial-temporal pandemic spread for SARS-CoV-2 (COVID-19) across large regions. In order to deal with the heterogeneity of COVID-19

spread in this scenario, it is proposed a novel SEIR-type model with Lagrangian-like movement where epidemic dynamics is described from behavior within and between time varying epidemic regions.

The methodology used for epidemic region identification overcome existing methodologies since local epidemic dynamic is captured at the same time that mobility relation between epidemic regions considers actual connections between geographical units. In addition, the design of algorithm for epidemic regions identification is translated to competence for identifying local outbreaks and reflecting the close relation between epidemic spread and socioeconomic population activity.

In regard numerical aspects, our methodology has few free parameters, all of them having epidemiological significance, and it is implemented using standard packages (namely Python ODE solver and Gephi software). That is, numerical procedure may be implemented with low computational cost in terms of software development and execution time.

The capabilities of our methodology are shown by analyzing the dynamics of the SARS-CoV-2 pandemic in the state of Jalisco, Mexico, as a case study. At the beginning of the study period (July 2020) the 125 municipalities of Jalisco were grouped into seven epidemic regions and by January 2021 nine regions are identified. The estimates for each epidemic region and for the state as a whole are satisfactory since the RRMSE obtained is below 15% in six from seven epidemic regions and equals to 4% for the whole Jalisco state. It should be noted that the epidemic regions maintain a good relationship with the 12 economic/geographical regions into which the 125 municipalities of Jalisco are officially grouped.

This methodology has been extended for analyze the epidemic dynamics for months from September, 2020, until February, 2021, obtaining similar results. Those results are shown as supplementary material.

During the development of this work, there were identified some issues that could complement it as future work. Incorporating a dynamic parameter estimation via data assimilation techniques would certainly improve this methodology. On the other hand, we are convinced that process for identifying epidemic regions could be reinforced by adding socioeconomic aspects: Indicators such as economic activity level, health care services access, education access level, infrastructure development, etc., whenever available, would be helpful and are obviously relevant for the epidemic dynamics. Without doubt, the NPI mechanisms modify the dynamic of contagions within municipalities. Analyzing effects of heterogeneous application of NPI policies is a key aspect that could be incorporated to ERSEIR model.

Finally, it is remarked that proposal methodology is extensible for modelling spread of diverse contagion agents and inclusive, to model socioeconomic dynamics.

## **Acknowledgments**

To the CONACYT for providing the postgraduate grant for R. Z. M. and the Universidad de Guadalajara for financial support through PROSNI program. Thanks to the anonymous reviewers for insightful comments that we believe have significantly improved this manuscript.

## **Conflict of interest**

The authors declare that they have no competing interests regarding the publication of this paper.

## References

1. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, *Lancet Infect. Dis.*, **20** (2020), 533–534.
2. *The World Bank*, World Bank open data, 2021. Available from: <https://databank.worldbank.org/home.aspx>.
3. SEGOB, Acuerdo por el que se establecen acciones extraordinarias para atender la emergencia sanitaria generada por el virus SARS-CoV2, 2020. Available from: [https://dof.gob.mx/nota\\_detalle.php?codigo=5590914&fecha=31%2F03%2F2020](https://dof.gob.mx/nota_detalle.php?codigo=5590914&fecha=31%2F03%2F2020).
4. INEGI, Censo de población y vivienda 2020, características de las localidades, tabulados predefinidos, 2021. Available from: <https://www.inegi.org.mx/programas/ccpv/2020/>.
5. I. Ibarra-Nava, J. A. Cardenas-De La Garza, R. E. Ruiz-Lozano, R. G. Salazar-Montalvo, Mexico and the COVID-19 response, *Disaster Med. Public Health Prep.*, **14** (2020), e17–e18. <https://doi.org/10.1017/dmp.2020.260>
6. F. Knaul, H. Arreola-Ornelas, T. Porteny, M. Touchton, M. Sánchez-Talanquer, Ó. Méndez, et al., Not far enough: Public health policies to combat COVID-19 in Mexico's states, *PLoS One*, **16** (2021), e0251722. <https://doi.org/10.1371/journal.pone.0251722>
7. M. Sánchez-Talanquer, E. González-Pier, J. Sepúlveda, L. Abascal-Miguel, J. Fieldhouse, C. Del Rio, et al., *Mexico's Response to COVID-19: A Case Study*, Institute for Global Health Sciences, University of California, San Francisco, 2021. Available from: <https://globalhealthsciences.ucsf.edu/sites/globalhealthsciences.ucsf.edu/files/mexico-covid-19-case-study-english.pdf>.
8. I. Franch-Pardo, B. M. Napoletano, F. Rosete-Verges, L. Billa, Spatial analysis and GIS in the study of COVID-19. A review, *Sci. Total Environ.*, **739** (2020), 140033. <https://doi.org/10.1016/j.scitotenv.2020.140033>
9. C. Scarpone, S. T. Brinkmann, T. Große, D. Sonnenwald, M. Fuchs, B. B. Walker, A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany, *Int. J. Health Geographics*, **19** (2020), 1–17. <https://doi.org/10.1186/s12942-020-00225-1>
10. N. Arinaminpathy, J. Das, T. H. McCormick, P. Mukhopadhyay, N. Sircar, Quantifying heterogeneity in SARS-CoV-2 transmission during the lockdown in India, *Epidemics*, **36** (2021), 100477. <https://doi.org/10.1101/2020.09.10.20190017>
11. Y. Ge, Z. Chen, A. Handel, L. Martinez, Q. Xiao, C. Li, et al., The impact of social distancing, contact tracing, and case isolation interventions to suppress the COVID-19 epidemic: A modeling study, *Epidemics*, **36** (2021), 100483. <https://doi.org/10.1016/j.epidem.2021.100483>
12. S. Triambak, D. P. Mahapatra, N. Mallick, R. Sahoo, A new logistic growth model applied to COVID-19 fatality data, *Epidemics*, (2021), 100515. <https://doi.org/10.1016/j.epidem.2021.100515>
13. P. Pincheira-Brown, A. Bentancor, Forecasting COVID-19 infections with the semi-unrestricted generalized growth model, *Epidemics*, **37** (2021), 100486. <https://doi.org/10.1016/j.epidem.2021.100486>

14. M. Fatima, K. J. O’Keefe, W. Wei, S. Arshad, O. Gruebner, Geospatial analysis of COVID-19: A scoping review, *Int. J. Environ. Res. Public Health*, **18** (2021), 2336. <https://doi.org/10.3390/ijerph18052336>
15. S. He, S. Tang, L. Rong, A discrete stochastic model of the COVID-19 outbreak: Forecast and control, *Math. Biosci. Eng.*, **17** (2020), 2792–2804. <https://doi.org/10.3934/mbe.2020153>
16. A. K. Srivastav, P. K. Tiwari, P. K. Srivastava, M. Ghosh, Y. Kang, A mathematical model for the impacts of face mask, hospitalization and quarantine on the dynamics of COVID-19 in India: deterministic vs. stochastic, *Math. Biosci. Eng.*, **18** (2021), 182–213. <https://doi.org/10.3934/mbe.2021010>
17. D. Olabode, J. Culp, A. Fisher, A. Tower, D. Hull-Nye, X. Wang, Deterministic and stochastic models for the epidemic dynamics of COVID-19 in Wuhan, China, *Math. Biosci. Eng.*, **18** (2021), 950–967. <https://doi.org/10.3934/mbe.2021050>
18. T. S. Brown, K. Engø-Monsen, M. V. Kiang, A. S. Mahmud, R. J. Maude, C. O. Buckee, The impact of mobility network properties on predicted epidemic dynamics in Dhaka and Bangkok, *Epidemics*, **35** (2021), 100441. <https://doi.org/10.1016/j.epidem.2021.100441>
19. W. Chiang, X. Liu, G. Mohler, Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates, *Int. J. Forecast.*, **38** (2021), 505–520. <https://doi.org/10.1016/j.ijforecast.2021.07.001>
20. R. Engbert, M. M. Rabe, R. Kliegl, S. Reich, Sequential data assimilation of the stochastic SEIR epidemic model for regional COVID-19 dynamics, *Bull. Math. Biol.*, **83** (2021), 1–16. <https://doi.org/10.1007/s11538-020-00834-8>
21. V. Gopalakrishnan, S. Pethe, S. Kefayati, S. Raman, P. Hake, A. Deshpande, et al., Globally local: Hyper-local modeling for accurate forecast of COVID-19, *Epidemics*, **37** (2021), 100510. <https://doi.org/10.1016/j.epidem.2021.100510>
22. R. Huang, M. Liu, Y. Ding, Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis, *J. Infect. Dev. Countries*, **14** (2020), 246–253. <https://doi.org/10.3855/jidc.12585>
23. Y. Mammeri, A reaction-diffusion system to better comprehend the unlockdown: Application of SEIR-type model with diffusion to the spatial spread of COVID-19 in France, *Comput. Math. Biophys.*, **8** (2020), 102–113. <https://doi.org/10.1515/cmb-2020-0104>
24. X. Hou, S. Gao, Q. Li, Y. Kang, N. Chen, K. Chen, et al., Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race, *Proc. Natl. Acad. Sci.*, **118** (2021). <https://doi.org/10.1073/pnas.2020524118>
25. T. W. Alleman, J. Vergeynst, L. De Visscher, M. Rollier, E. Torfs, I. Nopens, et al., Assessing the effects of non-pharmaceutical interventions on SARS-CoV-2 transmission in Belgium by means of an extended SEIQRD model and public mobility data, *Epidemics*, **37** (2021), 100505. <https://doi.org/10.1016/j.epidem.2021.100505>
26. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.*, **87** (2015), 925. <https://doi.org/10.1103/RevModPhys.87.925>

27. Y. Wang, J. Cao, Final size of network epidemic models: properties and connections, *Sci. China Inf. Sci.*, **64** (2021), 179201. <https://doi.org/10.1007/s11432-019-2656-2>
28. D. O'Sullivan, M. Gahegan, D. J. Exeter, B. Adams, Spatially explicit models for exploring COVID-19 lockdown strategies, *Trans. GIS*, **24** (2020), 967–1000. <https://doi.org/10.1111/tgis.12660>
29. D. Laroze, E. Neumayer, T. Plümper, COVID-19 does not stop at open borders: Spatial contagion among local authority districts during England's first wave, *Social Sci. Med.*, **270** (2021). <https://doi.org/10.1016/j.socscimed.2020.113655>
30. I. Franch-Pardo, M. R. Desjardins, I. Barea-Navarro, A. Cerdà, A review of GIS methodologies to analyze the dynamics of COVID-19 in the second half of 2020, *Trans. GIS*, **25** (2021), 2191–2239. <https://doi.org/10.1111/tgis.12792>
31. A. Gandolfi, Percolation methods for SEIR epidemics on graphs, in *Dynamic Models of Infectious Diseases*, Springer, New York, NY, (2013), 31–58. [https://doi.org/10.1007/978-1-4614-9224-5\\_2](https://doi.org/10.1007/978-1-4614-9224-5_2)
32. R. Wallace, L. Bergmann, L. Hogerwerf, R. Kock, R. G. Wallace, Ebola in the hog sector: Modeling pandemic emergence in commodity livestock, in *Neoliberal Ebola*, Springer, Cham, (2016), 13–53. [https://doi.org/10.1007/978-3-319-40940-5\\_2](https://doi.org/10.1007/978-3-319-40940-5_2)
33. R. Wallace, D. Wallace, H. Andrews, R. Fullilove, M. T. Fullilove, The spatiotemporal dynamics of AIDS and TB in the New York metropolitan region from a sociogeographic perspective: Understanding the linkages of central city and suburbs, *Environ. Plann. A*, **27** (1995), 1085–1108. <https://doi.org/10.1068/a271085>
34. U. N. Raghavan, R. Albert, S. Kumara, Near linear time algorithm to detect community structures in large-scale networks, *Phys. Rev. E*, **76** (2007), 036106. <https://doi.org/10.1103/PhysRevE.76.036106>
35. M. Martcheva, *An introduction to Mathematical Epidemiology*, Springer, **61** (2015). <https://doi.org/10.1007/978-1-4899-7612-3>
36. *Center for Disease Control and Prevention*, Scientific brief: SARS-CoV-2 transmission, 2022. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmission.html>.
37. J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, et al., Inferring the effectiveness of government interventions against COVID-19, *Science*, **371** (2021). <https://doi.org/10.1126/science.abd9338>
38. N. Imai, K. A. M. Gaythorpe, S. Abbott, S. Bhatia, S. van Elsland, K. Prem, et al., Adoption and impact of non-pharmaceutical interventions for COVID-19, *Wellcome Open Res.*, **5** (2020). <https://doi.org/10.12688/wellcomeopenres.15808.1>
39. N. L. Komarova, A. Azizi, D. Wodarz, Network models and the interpretation of prolonged infection plateaus in the COVID19 pandemic, *Epidemics*, **35** (2021), 100463. <https://doi.org/10.1016/j.epidem.2021.100463>
40. D. Pradhan, P. Biswasroy, P. Kumar Naik, G. Ghosh, G. Rath, A review of current interventions for COVID-19 prevention, *Arch. Med. Res.*, **51** (2020), 363–374. <https://doi.org/10.1016/j.arcmed.2020.04.020>

41. K. Soltesz, F. Gustafsson, T. Timpka, J. Jaldén, C. Jidling, A. Heimerson, et al., The effect of interventions on COVID-19, *Nature*, **588** (2020), E26–E28. <https://doi.org/10.1038/s41586-020-3025-y>
42. P. Bacchetti, N. P. Jewell, Nonparametric estimation of the incubation period of AIDS based on a prevalent cohort with unknown infection times, *Biometrics*, **47** (1991), 947–960. <https://doi.org/10.2307/2532651>
43. B. Cazelles, S. Hales, Infectious diseases, climate influences, and nonstationarity, *PLoS Med.*, **3** (2006), e328. <https://doi.org/10.1371/journal.pmed.0030328>
44. J. Lu, A. Lin, C. Jiang, A. Zhang, Z. Yang, Influence of transportation network on transmission heterogeneity of COVID-19 in China, *Transp. Res. Part C Emerging Technol.*, **129** (2021), 103231. <https://doi.org/10.1016/j.trc.2021.103231>
45. F. Brauer, P. V. den Driessche, J. Wu, *Mathematical Epidemiology*, Berlin, Germany, Springer, **1945** (2008). <https://doi.org/10.1007/978-3-540-78911-6>
46. J. Chen, S. Levin, S. Eubank, H. Mortveit, S. Venkatramanan, A. Vullikanti, et al., Networked epidemiology for COVID-19, *SIAM News*, **53** (2020). Available from: <https://par.nsf.gov/servlets/purl/10213734>.
47. M. Marathe, A. K. S. Vullikanti, Computational epidemiology, *Commun. ACM*, **56** (2013), 88–96. <https://doi.org/10.1145/2483852.2483871>
48. W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. London, Ser. A*, **115** (1927), 700–721. <https://doi.org/10.1098/rspa.1927.0118>
49. F. Arrigoni, A. Pugliese, Limits of a multi-patch SIS epidemic model, *J. Math. Biol.*, **45** (2002), 419–440. <https://doi.org/10.1007/s002850200155>
50. C. Wolf, M. Langlais, D. Ponter, A multi-patch epidemic model with periodic demography, direct and indirect transmission and variable maturation rate, *Math. Popul. Stud.*, **13** (2006), 153–177. <https://doi.org/10.1080/08898480600788584>
51. J. Arino, J. M. Miller, P. Van den Driessche, A multi-species epidemic model with spatial dynamics, *Math. Med. Biol.*, **22** (2005), 129–142. <https://doi.org/10.1093/imammb/dqi003>
52. R. K. McCormack, L. J. S. Allen, Multi-patch deterministic and stochastic models for wildlife diseases, *J. Biol. Dyn.*, **1** (2007), 63–85. <https://doi.org/10.1080/17513750601032711>
53. D. Bichara, A. Iggidr, Multi-patch and multi-group epidemic models: a new framework, *J. Math. Biol.*, **77** (2018), 107–134. <https://doi.org/10.1007/s00285-017-1191-9>
54. M. Girvan, M. E. J. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci.*, **99** (2002), 7821–7826. <https://doi.org/10.1073/pnas.122653799>
55. F. Parés, D. G. Gasulla, A. Vilalta, J. Moreno, E. Ayguadé, J. Labarta, et al., Fluid communities: A competitive, scalable and diverse community detection algorithm, in *International Conference on Complex Networks and Their Applications*, Springer, Cham, (2017), 229–240. [https://doi.org/10.1007/978-3-319-72150-7\\_19](https://doi.org/10.1007/978-3-319-72150-7_19)
56. S. Fortunato, Community detection in graphs, *Phys. Rep.*, **486** (2010), 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>

57. A. Clauset, M. E. J. Newman, C. Moore, Finding community structure in very large networks, *Phys. Rev. E*, **70** (2004), 066111. <https://doi.org/10.1103/PhysRevE.70.066111>
58. M. Bastian, S. Heymann, M. Jacomy, Gephi: An open source software for exploring and manipulating networks, in *Proceedings of the Third International Conference on Weblogs and Social Media*, 2009. <https://doi.org/10.13140/2.1.1341.1520>
59. R. Wallace, Y. S. Huang, P. Gould, D. Wallace, The hierarchical diffusion of AIDS and violent crime among US metropolitan regions: inner-city decay, stochastic resonance and reversal of the mortality transition, *Social Sci. Med.*, **44** (1997), 935–947. [https://doi.org/10.1016/s0277-9536\(96\)00197-9](https://doi.org/10.1016/s0277-9536(96)00197-9)
60. Secretaría de Salud Jalisco, Datos de Radar Jalisco, 2021. Available from: <https://coronavirus.jalisco.gob.mx/datos-abiertos>.
61. H. Gutiérrez-Pulido, D. Gómez-Alvarez, Estudio técnico para la adecuación de las regiones de Jalisco, Technical report, Gobierno de Jalisco, 2016. Available from: <https://seplan.app.jalisco.gob.mx/biblioteca/archivo/verDocumento/939>.
62. K. Ejima, K. S. Kim, C. Ludema, A. I. Bento, S. Iwanami, Y. Fujita, et al., Estimation of the incubation period of COVID-19 using viral load data, *Epidemics*, **35** (2021), 100454. <https://doi.org/10.1016/j.epidem.2021.100454>
63. H. Najafimehr, K. M. Ali, S. Safari, M. Yousefifard, M. Hosseini, Estimation of basic reproduction number for COVID-19 and the reasons for its differences, *Int. J. Clin. Pract.*, **74** (2020), e13518. <https://doi.org/10.1111/ijcp.13518>
64. W. He, G. Y. Yi, Y. Zhu, Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis, *J. Med. Virol.*, **92** (2020), 2543–2550. <https://doi.org/10.1002/jmv.26041>
65. S. Zhao, B. Tang, S. S. Musa, S. Ma, J. Zhang, M. Zeng, et al., Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data, *Epidemics*, **36** (2021), 100482. <https://doi.org/10.1016/j.epidem.2021.100482>
66. J. A. Backer, D. Klinkenberg, J. Wallinga, Incubation period of 2019 novel coronavirus (2019-NCOV) infections among travellers from Wuhan, China, 20–28 January 2020, *Eurosurveillance*, **25** (2020), 2000062. <https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062>
67. R. J. Hyndman, A. B. Koehler, Another look at measures of forecast accuracy, *Int. J. Forecast.*, **22** (2006), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>



©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)