Gravity Model Approach to Model Epidemic with Human Dispersal Behaviors

A.S.K. Dinasiri[1*], Y. Jayathunga[1]

[1]Department of Mathematics, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka

[*]Email: askd99sl@gmail.com

Abstract

The gravity model which is based on Newton's gravitational law, has been widely used as a spatial interaction model in the past few decades. Spatial interactions are important in epidemic modeling as different populations in the world are interconnected by them. Human dispersal behaviors are spatial interactions and they are crucial aspects of infectious disease spread. However, many existing compartmental models model epidemics in a single area. Hence, a gravity model approach to model epidemics incorporated with a multi-patch compartmental model is studied here. Both human dispersal behaviors within a patch and between patches are considered. When the human dispersal behaviors within a patch are modeled, the denominator of the general gravity model becomes zero. An alternative power-based distance decay function is introduced to the gravity model to address that research gap. The parameters of the modified gravity model are estimated using a hybrid method combining ordinary least squares (OLS) and nonlinear least squares (NLS) methods (Hybrid OLS-NLS method).

Keywords: Gravity model with human dispersal behaviors, distance decay function, multi-patch compartmental model, OLS, NLS, hybrid OLS-NLS method

2020 MSC classification number: 92D30, 93A30

1. INTRODUCTION

In the world, a sudden unexpected increase in the number of infected humans with an infectious disease can be experienced at any time. According to the Centers for Disease Control and Prevention, such an increment in the number of infected humans, which occurs in a specific geographical area is considered as an epidemic [1]. From the Plague of Athens during 430 - 426 B.C.E to COVID-19 since 2019, human society has been replete with various kinds of epidemics such as the Black Death, Cholera, Russian Flu, MERS, etc [2], [3]. Those occurrences of epidemics make people motivated continuously to study epidemics.

An epidemic occurs in a sudden disease outbreak. A large number of people in a particular population, region, or community, is affected by an epidemic [4]. Epidemic modeling plays a significant role in providing insight into the transmission dynamics of communicable diseases [5], [6], [7]. The spread of those diseases through a community can be described from such epidemic models [8]. Those models consist of a set of equations that translate the assumptions into a mathematical representation [9].

Human dispersal behaviors play a major role in infectious disease spread. Due to globalization and transportation development, an increase in human dispersal behavior can be observed. Due to the increment of human dispersal behaviors, the transmission speed of infectious diseases got faster [10].

The classical compartmental models are not feasible for incorporating the disease spread with spatial interactions [11]. A more accurate model can be obtained by concentrating on spatially distributed sub-populations [12]. Since the gravity model which is based on Newton's gravitational laws is a spatial interaction model, it is motivated to consider a multi-patch model with the gravity model approach incorporating a compartmental model [13]. The classical epidemic models are mostly based on the assumption that the entire population is well-mixed, and they live in one area. But usually, humans live in different areas and the disease spread is affected by spatial heterogeneity [14]. In epidemiology, such areas are called patches. Meta-populations occur if those patches are connected through human dispersal behaviors [2]. Human dispersal behaviors play a major role in the COVID-19 spread as it is an infectious disease [15].

*Corresponding author

Received May 1^{st}, 2024, Revised October 20^{th}, 2024, Accepted for publication December 11^{th}, 2024. e-ISSN: 2549-2896, DOI:10.5614/cbms.2024.7.2.5

In the early stages of COVID-19, traveling restrictions and lockdowns were some of the pivotal non-pharmaceutical interventions (NPIs) used to limit human dispersal behaviors and NPIs dropped off the speed of the spread [16]. Several studies have been conducted to model epidemics with human dispersal behaviors. A study that is focused on understanding the impact of mobility on COVID-19 spread presents the relationship between the spatial interactions and the distance between patches. The spatial interaction is inversely proportional to the distance [16].

According to Newton's law of gravity, the gravitational force between two masses is directly proportional to the product of the two masses and it is inversely proportional to the square of the distance between them. The typical gravity model is based on the Newton's law of gravity. It can be written as:

$$F = \frac{Gm_1m_2}{r^2}, \tag{1}$$

where F is the gravitational force, m_1 and m_2 are the masses of two objects, r is the distance between them and G is a constant [17]. Steward, who was a Princeton astronomer, noticed that Newton's law of gravity is resembled by the spatial interactions. Then the gravity law is incorporated with the spatial interactions by Steward during 1940s. He expressed it as:

$$F = \frac{GP_iP_j}{d_{ij}^2}, \tag{2}$$

where F is the demographic force, P_i and P_j are the populations of the origin and the destination respectively, d_{ij} is the distance between them and G is a constant [18].

Various forms of gravity models inspired by Newton's law of gravity are widely used in different areas which are incorporated with the spatial interactions. The gravity model has been used in the area of international trade, traffic flows, and human mobility [19], [20]. In 1962, Tinbergen introduced the gravity model concepts into international trade [21]. The gravity model was used for migration flows by Ravenstein in 1889 [22].

The gravity model is developed for human dispersal behaviors. It says that the demographic force is directly proportional to the size of the populations in the patches and inversely proportional to the distance between them [18]. Then the gravity model is written as follows:

$$F_{ij} = K\frac{{P_i}^n{P_j}^m}{d_{ij}^\gamma}. \tag{3}$$

Since the gravity model is adapted in various domains, generally it can be written as:

$$F_{ij} = K{P_i}^n{P_j}^m\mathrm{f}_\gamma(d_{ij}). \tag{4}$$

With the various studies which have been conducted, the form of a gravity model for the human dispersal behaviors has been generalized and modified. It is influenced by a form of distance decay function $(\mathrm{f}_\gamma(d_{ij}))$ where the spatial and social interactions increase as the distance between two locations i and j (d_{ij}) decreases [16]. The distance decay function can be a power law-based function or an exponential law-based function. However, it depends on the distance between the two locations [23].

Table 1: The description of the gravity model used for calculating human dispersal behaviors.

Symbol	Description
F_{ij}	Number of people moving from i^{th} patch to j^{th} patch
P_i, P_j	Population in i^{th} and j^{th} patches
d_{ij}	Geographical distance between patch i and j
K	Proportionality constant
n, m, γ	Parameters to be estimated
N	Number of patches.

To model a compartmental model with a gravity model approach, creating a mobility matrix is a major concern. If someone uses the gravity model for human dispersal behaviors within a patch, F_{ij} becomes undefined because the distance from a patch to the same patch d_{ij} is zero. There was a lot of literature

that tried to overcome that problem. In literature, the gravity model was used to find the number of trips between two geographical locations. The probability of mobility within a patch was explained with weighted average methods [24]. In some literature, the volume of interactions (contacts) between patches is given by a gravity model approach. There they used the typical gravity equation in Equation 2 for modeling the contacts between two patches. The radius of a circle with an area equal to the area of the patch is used to d_{ij} instead of the distance to model the contacts within the patch [12].

The researchers are interested in using a gravity model approach to model an epidemic in a patchy environment. A study on a two-patch system brings out the gravity model to model the rate of change of the number of humans in a particular patch. That does not model human mobility within a patch [25]. The rate of mobility within and between patches is modeled separately in previous incidences. Therefore, it is interesting to look for a method that can incorporate the rate of human dispersal behaviors within and between patches in the same gravity model approach. The gravity model is used as a model to measure human dispersal behaviors. In a study on the influenza epidemic in the United States of America, the major aspects of workflows are apprehended using a piece-wise gravity model [26].

The infectious diseases are classified as human-to-human transmitted diseases, food and waterborne diseases, and vector-borne diseases [2]. Since the study is directed to a gravity model approach with human dispersal behavior, the study is focused on human-to-human transmitted diseases.The recent infectious disease; COVID-19 which later became a global epidemic is a viral infectious disease that emerged in Wuhan City, Hubei province, China in December 2019 [27]. It was caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). It transmits from human to human by direct contact and droplets. It was spread all over the world with a $2\% - 3\%$ fatality rate [28], and resulted in 775,132,086 cases including 7,042,222 deaths globally [29]. Including Sri Langka, 229 countries and territories were burdened by COVID-19. In Sri Lanka, the first patient found on the 27^{th} of January 2020 was a Chinese woman, and the first confirmed local case was reported on the 11^{th} of March 2020 in Sri Lanka [30]. Then the whole population was susceptible and 672 754 confirmed cases with 16 889 deaths were reported [31].

Sri Lanka is divided into 9 provinces for administrative purposes which are named Western, Southern, Central, North Western, North Central, Uva, Sabaragamuwa, Northern, and Eastern. Human dispersal behaviors spread through those provinces due to educational purposes, career purposes, freight transportation, and various other reasons.

In Sri Lanka, the gravity model approach is used to study the apparel exports, exports, trade, and decision support areas [32], [33]. There is room to research about the transmission of COVID-19 from a Sri Lankan perspective by a gravity model approach with human dispersal behaviors. To address these gaps, a gravity model approach to model human-to-human transmitted epidemic with human dispersal behaviors with a specific focus on studying the spread of COVID-19 within and between different provinces with human dispersal behaviors in Sri Lanka can be started.

The main objective of this research is to formulate a multi-patch epidemic model via a gravity model approach by identifying a way to model human dispersal behaviors within a patch using the gravity model. It incorporates the probability matrix of human dispersal behaviors by using the gravity model concept to investigate the transmission patterns of an epidemic numerically with human dispersal behaviors.

2. Model formulation

2.1. Modified gravity model

The gravity model discussed in Equation 4 is used to create the mobility matrix for human dispersal behaviors. Since the generalized gravity model discussed in Equation 3 can not express the mobility within a patch, a need to find a way to express the mobility within a patch is raised. The distance from a patch to the same patch is zero. The mobility will be undefined if the gravity model in Equation 3 is used. Because a division of zero occurs then. Hence, a suitable function for the distance decay function $(\mathrm{f}_\gamma(d_{ij}))$ mentioned in Equation 4 must be chosen.

The gravity models can be studied in two ways concerning the type of distance decay functions. The power-based gravity model and the exponential-based gravity models are the two types. In the power-based gravity model, an inverse power function is used as the distance decay function. A negative exponential function is used as the distance decay function for the exponential-based gravity model. In Newton's gravity model, the distance decay function is an inverse power function.

An action at a distance that properly reflects the spatial interaction nature is suggested by the inverse power function. But a reflection of spatial interactions between distant patches can not be gained by the negative exponential function [34].

There are two laws of Geography stated as follows by Waldo R. Tobler and G. Arbia.

1) "Everything is related to everything else, but near things are more related than distant things [35]."
2) "Everything is related to everything else, but things observed at a coarse spatial resolution are more related than things observed at a finer resolution [36]."

These laws are based on the typical gravity model and action at a distance. The inverse proportional property of the distance and the direct proportional property of the sizes of populations are reflected by the above laws respectively. But the negative exponential distance decay function refuses it. Also power based gravity model can express complex networks. Hence, choosing the inverse power function as the distance decay function is more appropriate than choosing the negative exponential distance decay function [34].

In a study for understanding the impact of human dispersal behaviors on the spread of infectious disease, the human dispersal behaviors within and between patches are captured based on the relationship between distance. A model is presented as an inverse power distance decay function by considering that relationship as,

$$\mathrm{f}_\gamma(d_{ij}) = \frac{1}{(1+d_{ij})^\gamma}, \tag{5}$$

where d_{ij} is the distance between patch i and j and γ is a parameter to be estimated [16]. This solves the problematic scenario of division by zero that occurs in the gravity model discussed in Equation 3. Hence, the modified gravity model can be written as follows:

$$F_{ij} = K\frac{{P_i}^n {P_j}^m}{(1+d_{ij})^\gamma}. \tag{6}$$

All the symbols that are used in Equation 6 are described in Table 1.

2.2. Parameter estimation

*1) **Using OLS method**:* The typical gravity model explained in Equation 1, has its' population exponents as 1 and the distance exponent as 2. But in the context of epidemiology, the gravity model concept is used to capture the human dispersal behaviors. According to many studies, the exponents in the modified gravity model in Equation 6 rarely be the same as the typical gravity model in Equation 1 [23], [25], [26]. Hence, a parameter estimation must be executed to find the appropriate values for the parameters n, m, γ, and K in the modified gravity model (Equation 6).

The most common way to estimate the gravity parameters is the ordinary least squares (OLS) method. To perform the OLS estimation, the modified gravity model equation must be log-linearised as follows:

$$\ln F_{ij} = \ln K + n\ln P_i + m\ln P_j - \gamma\ln(1+d_{ij}). \tag{7}$$

Then the mobility data, population data, and the distances related to the nine provinces in Sri Lanka are used to fit the log-linearised model in Equation 7. The built-in function `'minimize'` in `scipy.optimize` in `SciPy` python library, is used to minimize the sum of squares of the residual sum. Here initial guesses for the parameters n, m, γ and K must be provided to complete the procedure.

*2) **Initial guesses**:* The initial guesses for n, m, and γ are guessed from literature. A study of disease spread with a piecewise gravity model that incorporates workflows gives two sets of values for n, m, and γ such as $(0.30, 0.64, 3.05)$ and $(0.24, 0.14, 0.29)$ [26]. Also, γ is found to be in the range of $(0.5 - 3.0)$ in various studies for different data sets [20]. Hence, the initial guesses are chosen as $n = 0.27, m = 0.39$, and $\gamma = 1.67$.

To make an initial guess for the proportionality constant K, the following steps that are widely used in the area of physics are taken into consideration. First, the modified gravity model in Equation 6 must be written

in the form of $y = m_0 x$ where F_{ij} acts as the dependent variable (y) and $\frac{{P_i}^n {P_j}^m}{(1+d_{ij})^\gamma}$ acts as the independent variable (x). Then the K starts to work as the slope (m_0).

$$\underbrace{F_{ij}}_{y} = \underbrace{K}_{m_0} \underbrace{\frac{{P_i}^n {P_j}^m}{(1+d_{ij})^\gamma}}_{x} \tag{8}$$

Here the parameter values used in sub-section 2.2.2 are used to evaluate the x values in Equation 8. To find the value of K, a polynomial with degree 1 which is a line is fitted to the data. The built-in function `'polyfit()'` in Python `NumPy` library is used to do that task. Then $\ln K$ is given as the initial guess to the OLS method where $\ln K = 4.83$. Therefore, $(0.27, 0.39, 1.67, 4.83)$ are given as the initial guesses for $(n, m, \gamma, \ln K)$ in OLS method.

*3) **Using NLS method**:* The OLS method for a log linearised model has some drawbacks in estimating parameters. It can not perform parameter estimation if the dependent variable of the model becomes zero. In the context of our study, parameter estimation could not be done by OLS if there is zero mobility from patch i to patch j.(i.e. $F_{ij} = 0$). Because when we log-linearise the model, then $\ln F_{ij}$ becomes undefined as $\ln 0$ has not been defined. This problem is not encountered with the nonlinear least squares (NLS) method. Also, it can estimate parameters without log linearising the model. The built-in function `'least_squares'` in `scipy.optimize` in `SciPy` python library, is used to minimize the sum of squares of the residual sum. Here also initial guesses for the parameters n, m, γ, and K must be provided to complete the procedure.

*4) **Using the Hybrid OLS-NLS method**:* The NLS method is sensitive to the initial guesses. Therefore, it is important to have proper initial guesses for the parameter values to be estimated before using the NLS method.

Considering all aspects, the hybrid OLS-NLS method is introduced to estimate parameters by combining OLS and NLS methods. The steps are discussed below.

1) Choose appropriate initial guesses for the parameters to be estimated as discussed in sub-section 2.2.2.
2) Log linearise the modified gravity model as Equation 6.
3) Estimate parameters using the OLS with above mentioned initial guesses.
4) Use the estimated parameters from the OLS, as the initial guesses for the NLS method.
5) Estimate the parameters using the NLS method.

The estimated parameters will be discussed in Section 3.1

2.3. Mobility matrix

In this stage, the parameters in the modified gravity model 6 have been estimated as relevant to the mobility data of the provinces in Sri Lanka. The gravity model gives the number of people traveling from patch i to patch j. The mobility rate from patch i to j is denoted as p_{ij}. It is calculated by considering the proportion of patch i humans who are traveling in patch j. And it gives the probability of moving a human from patch i to patch j. Also, the mobility matrix which is the probability matrix of human dispersal behaviors can be denoted as,

$$\pi = (\pi_{ij})_{9\times 9}, \quad i, j \quad = 1, 2, ..., 9, \tag{9}$$

where i and j are the provinces.

Then the mobility rate from province i to j can be determined by,

$$\pi_{ij} = \frac{F_{ij}}{\sum_{j=1}^{9} F_{ij}}, \tag{10}$$

where $i, j = 1, 2, ..., 9$. Therefore, $\sum_{j=1}^{9} \pi_{ij} = 1$. Also the diagonal entries of the mobility matrix $(\pi_{ii}), \forall i \in \{1, 2, ..., 9\}$ provide the probability of moving a human from his residence patch to the same. The calculated mobility matrix which gives the probability matrix of human dispersal behaviors within and between the 9 provinces in Sri Lanka will be discussed in the Section 3.1.

2.4. Residence budgeting time matrix

The residence budgeting time matrix for the nine provinces in Sri Lanka is linked with the mobility matrix for nine provinces which is given in Equation 9. Most humans in Sri Lanka make short-term movements within and between provinces because of education and career purposes. Generally, Sri Lankan people have 8 working hours and we assume that it takes 2 hours to travel from home to the workplace and back to home. Therefore, we assume the proportion of time that a human travels from province i to other provinces (t_i) is $\frac{10}{24}$ Then,

$$P = \frac{14}{24} I_9 + \frac{10}{24}\pi, \tag{11}$$

where $\pi = (\pi_{ij})_{9\times 9}$, $P = (p_{ij})_{9\times 9}$, and I_9 is the identity matrix of size 9. Also, $\sum_{j=1}^{9} \pi_{ij} = \sum_{j=1}^{9} p_{ij} = 1$, $\forall i \in \{1, 2, ..., 9\}$.

2.5. Multi-patch SIR model with human dispersal behaviors

The classical SIR model is formulated on a well-mixed population that lives in the same area. However, in real-world scenarios, the spatial heterogeneity must be included in the model. The spatial interactions and the connection between different populations in patches occur with the human dispersal behaviors [2]. The movements of individuals that are caused by human dispersal behaviors can be classified as short-term movements and long-term movements.

When humans spend more time in their residence patch and visit another patch or patches to spend less time in that patch, such a movement is called a short-term movement. They come back to their residence patch after spending a certain proportion of time in another patch. Humans settle in another patch which is not their residence patch and those movements are called long-term movements. The modeling approaches are different concerning the type of dispersal behavior. The Lagrangian approach which is related to fluid dynamics is used when it comes to short-term movements and it becomes the Eularien approach with long-term movements [37]. The study highlights the Lagrangian approach.

First, we assume that the total population consists of N number of patches that are connected with human dispersal behaviors. Those patches can be geographical locations or any groups of people. We assumed that any individual in a patch visits another patch to spend some time and return to the residence patch. An individual can infect another individual in his residence patch or any other patch. Also, a healthy individual can be infected at his residence patch or any other patch [2].

In each patch, there are three compartments which are susceptible, infected, and recovered classes. The summation of those three compartments gives the total number of humans in a patch. $N_i = S_i + I_i + R_i$. The total population N_{total} is given by, $N_{total} = \sum_{i=1}^{N} N_i$.

Table 2: The description of the symbols used in multi-patch SIR model.

Symbol	Description
S_i	The number of susceptible individuals in patch i
I_i	The number of infected individuals in patch i
R_i	The number of recovered individuals in patch i
N_i	The total population in patch i
N	Number of patches
p_{ij}	The proportion of time that humans in patch i move in patch j
β_j	Transmission rate in patch i
α_i	Recovery rate in patch i

The multi-patch (meta-population) model represents its model in a large system of ordinary differential equations [2]. The ODE system can be written as follows [38]:

$$\begin{aligned}
\frac{dS_i}{dt} &= -S_i \sum_{j=1}^{N} \beta_j p_{ij} \left(\frac{\sum_{k=1}^{N} p_{kj} I_k}{\sum_{k=1}^{N} p_{kj} N_k} \right), && S_i(t_0) \geq 0, \\
\frac{dI_i}{dt} &= S_i \sum_{j=1}^{N} \beta_j p_{ij} \left(\frac{\sum_{k=1}^{N} p_{kj} I_k}{\sum_{k=1}^{N} p_{kj} N_k} \right) - \alpha_i I_i, && I_i(t_0) \geq 0, \\
\frac{dR_i}{dt} &= \alpha_i I_i, && R_i(t_0) \geq 0,
\end{aligned} \tag{12}$$

where $i, j, k = 1, 2, ..., N$ and $\sum_{j=1}^{N} p_{ij} = 1$.

In Sri Lanka, there are nine distinct provinces for administrative purposes [39]. They can be considered as patches that are connected with human dispersal behaviors. A type of multi-patch compartmental model discussed in the Equation 12, which is a nine-patch SIR model is used to investigate the COVID-19 spread within and between Sri Lanka. The results calculated by the residence budgeting time matrix are used for the p_{ij} values in Equation 12.

The transmission rate β_j, captures the risk of infection in patch j. And the infection rate of the patch i is,

$$\sum_{j=1}^{9} \beta_j p_{ij} \left(\frac{\sum_{k=1}^{9} p_{kj} I_k}{\sum_{k=1}^{9} p_{kj} N_k} \right).$$

Here the number of newly infected humans belonging to patch i who got infected while moving within their residence patch i is given by,

$$S_i \beta_i p_{ii} \left(\frac{\sum_{k=1}^{9} p_{ki} I_k}{\sum_{k=1}^{9} p_{ki} N_k} \right).$$

The number of new infected humans belongs to patch i who got infected while moving between their residence patch i and the other patches j is given by,

$$S_i \sum_{j=1, j\neq i}^{9} \beta_j p_{ij} \left(\frac{\sum_{k=1}^{9} p_{kj} I_k}{\sum_{k=1}^{9} p_{kj} N_k} \right).$$

Also, the proportion of time that residents from patch j spend in patch j is denoted by p_{jj} and p_{ij} denotes the proportion of time that residents from patch i spend in patch j.

Here all the infected humans are assumed to be infectious and all the recovered humans are assumed not to be susceptible again. Since the time duration of the disease is negligible compared to the lifetime of a human, the total population in the country is assumed to be constant.

2.6. Basic reproduction number ($\mathcal{R}_0$)

The next generation matrix method which is introduced by Diekmann, is widely used to derive the reproduction number for models that incorporate the spatial structures [40]. The next generation matrix G is created as, $G = FV^{-1}$, where F and V are the jacobians of the new infections appearance rate in the i^{th} compartment ($\mathcal{F}_i(\vec{\mathbf{x}})$) and the difference of the transfer rate of individuals into the compartment i and out of the compartment i ($\mathcal{V}_i(\vec{\mathbf{x}})$) at the disease free equilibrium point.

$$F = \frac{\partial \mathcal{F}_i(x_0)}{\partial x_j} \quad \text{and} \quad V = \frac{\partial \mathcal{V}_i(x_0)}{\partial x_j}.$$

Here the vector $\vec{\mathbf{x}}$ is denoted as, $\vec{\mathbf{x}} = x_i, i = 1, 2, .., q$ where q is the number of compartments in the system and x_i is the size of the i^{th} compartment or the proportion of individuals in the i^{th} compartment. Then the dominant eigenvalue of matrix G which is the spectral radius of G is defined as the reproduction number ($\mathcal{R}_0$) [41].

3. Numerical Results

3.1. Parameter estimation of the gravity model

The parameter estimation for the modified gravity model in Equation 6 is performed using OLS, NLS, and hybrid OLS-NLS methods as discussed in Section 2.2. The values described in sub-section 2.2.2 are given as the initial guesses for the estimation. The following are the estimated values for the gravity model parameters relevant to each estimation method.

Table 3: The estimated parameter values for the gravity model using OLS, NLS, and hybrid OLS-NLS methods.

	OLS method	NLS method	Hybrid OLS-NLS method
n	0.69994	0.27105	0.38326
m	0.94025	0.39114	0.60072
γ	0.83117	1.66548	0.77417
K	1.00007	124.72659	1.00977

From the three types of above estimation methods, three sets of parameters are obtained as the gravity model parameters. The following set of parameters which is obtained by the hybrid OLS-NLS method, is used as the parameters of the modified gravity model in Equation 6 in this research. The reason behind choosing this set of parameters is discussed in Section 4 considering the model evaluation metrics.

$$n = 0.38326, \quad m \quad = 0.60072, \quad \gamma = 0.77417, \quad K \quad = 1.00977 \tag{13}$$

3.2. The mobility matrix

The gravity model approach is used to create the mobility matrix by following the method discussed in Section 2.3. The distances between the nine provinces, populations in the nine provinces, and the parameter values estimated in Equations 13 are provided to the modified gravity model in Equation 6. Then the mobility matrix π is created as the above Figure 1. It shows that the probability within a province takes a higher value than between provinces.

Generally, the mobility within a patch takes a higher value than between the patches. That aspect is also captured by this gravity model approach and represented in diagonal values $(\pi_{ii}), \forall i \in \{1, 2, ..., 9\}$ of the mobility matrix π.

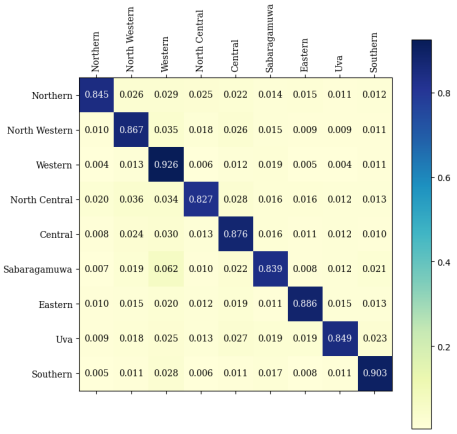

Figure 1: The mobility matrix for nine provinces in Sri Lanka obtained by the gravity model approach.

3.3. Visualization of the probabilities between provinces

Here we visualize the pattern of mobility between provinces. Mobility within a province is not included here.

*1) **Visualization with distances**:* Figure 2 shows how the human dispersal behaviors between provinces behave concerning the distance between the provinces. The values of the x-axis of each plot in Figure 2 are sorted in ascending order concerning the distance from each province to the other provinces.

The probabilities of human dispersal behaviors are almost reduced with the increment of the distance between provinces according to Figure 2. However, an increment is observed in the Western province in each plot though some other provinces have higher distances than the Western Province. According to the Equation 6, the populations of the provinces and the distance between them affect the human dispersal behaviors. The Western province has the highest population and it is the reason for the above-mentioned observation.

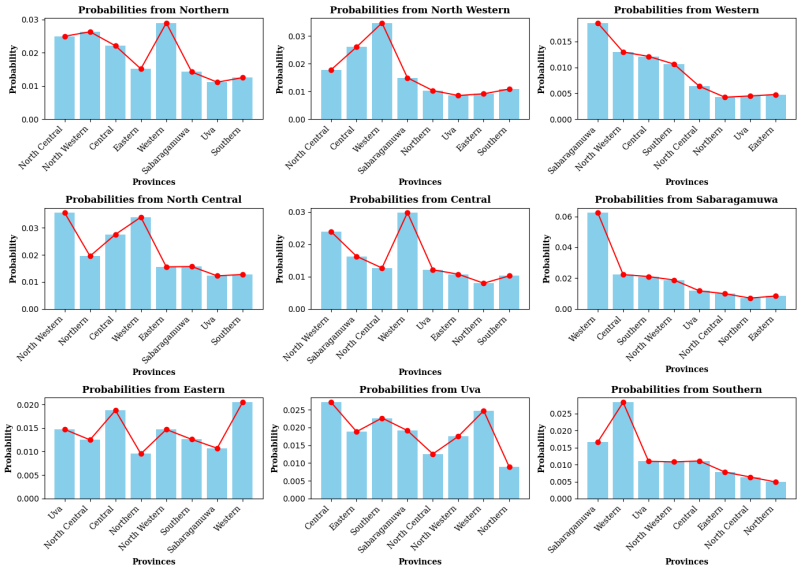

Figure 2: Human dispersal behaviors with the distances between provinces in each province in Sri Lanka.

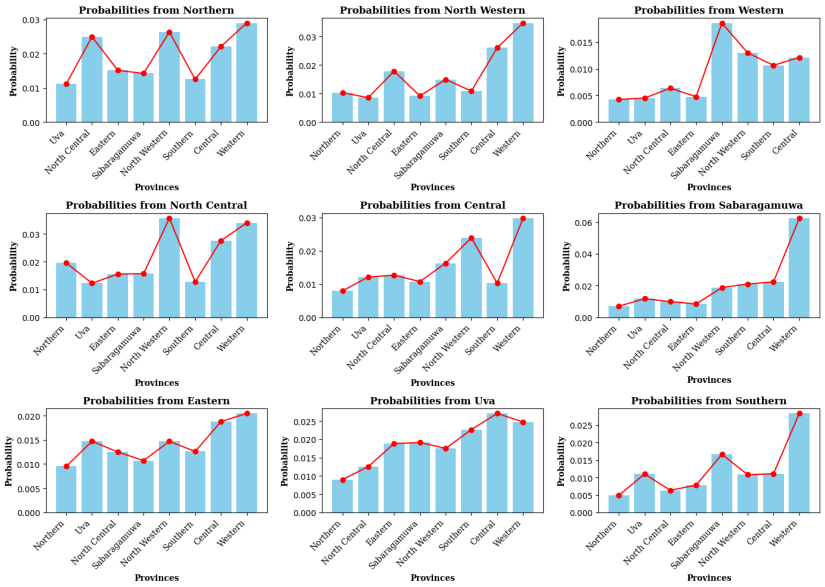

Figure 3: Human dispersal behaviors between provinces with the population size of the destination province in each province in Sri Lanka.

*2) **Visualization with population**:* The behavior of the human dispersal behaviors between provinces concerning the population of the destination province for each province is shown in Figure 3. The x-axis values of each plot in Figure 3 are sorted in ascending order concerning the population of the destination province. The probabilities of human dispersal behaviors are increasing when the population of the destination province is high. But it is changed due to the effect from distance as discussed through the modified gravity model in Equation 6.

*3) **Visualization with distance and population**:* The effects of both distance and population on the probability and the human dispersal behavior between provinces are presented in Figure 4. It represents the relationship between the probability of human dispersal behaviors between provinces with the population of the destination province and the distances between provinces in each province in Sri Lanka. The values of the x-axis and the values of the y-axis are sorted in ascending order concerning the distance between the provinces and the population of the destination province respectively.

3.4. The residence budgeting time matrix

Assuming that a human in i^{th} province spends 10 hours in any other province, the residence budgeting time matrix can be obtained using Equation 11.

3.5. Parameter estimation of multi-patch SIR model

The multi-patch SIR model in Equation 12, is used to incorporate the human dispersal behaviors into the COVID-19 disease spread. The calculated number of infectious humans using Equation 12 and the real data for infectious humans are considered to estimate the transmission rates for each province in Sri Lanka as shown in Table 4.

*1) **Fixed parameters**:* Sri Lanka is a small island with a small crowd of population compared to some other big countries like India, the USA, Russia, etc. Therefore, the recovery rates (α_i) of the humans in province i for all nine provinces are assumed to be the same α. Therefore, the recovery rate α for all provinces in Sri Lanka is considered as a fixed parameter such that $\alpha^{-1} = 10.25$ days [42].

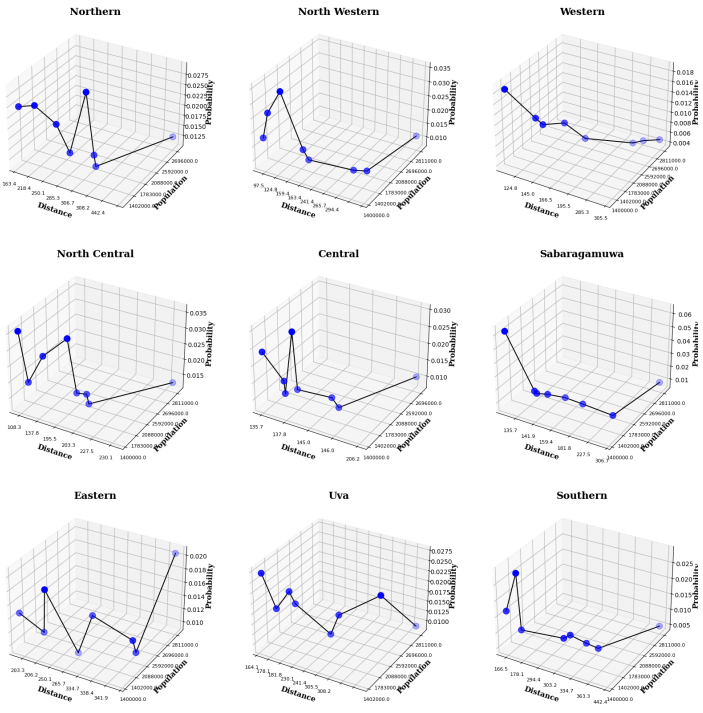

Figure 4: Human dispersal behaviors with the population of the destination province and the distances between provinces.

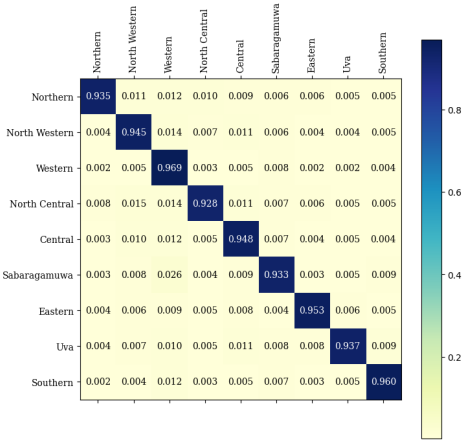

Figure 5: The residence budgeting time matrix for nine provinces in Sri Lanka obtained by the gravity model approach.

2) ***Estimated parameters****:* The estimated parameter values for the β_i for each province are shown in the Table 4. Neither lockdown nor social distancing strategies are considered in the estimation. Those estimations are done as the methods discussed in follows.

The built-in function `'least_squares'` in `scipy.optimize` in `SciPy` python library, is used to estimate the β_i values. The non-linear least squares method is used to minimize the sum of squares of the residual sum. The residuals are the difference between the actual data and the predicted data from the relevant ODE system.

The data on the number of infected and susceptible humans, populations in each province, and the residence budgeting time matrix time obtained in 5 by the gravity model approach is used in the estimating process. Also, the recovery rates for each province (α_i) are considered to have the same value α. For this research, Northern, North Western, Western, North Central, Central, Sabaraganuwa, Eastern, Uva, and Southern provinces are indexed as $1, 2, 3, 4, 5, 6, 7, 8,$ and 9 respectively.

Table 4: Estimated values for the transmission rates (β_i) of each province.

Symbol	Related Province	Value
β_1	Northern province	0.07344980
β_2	North Western province	0.13890953
β_3	Western province	0.12895880
β_4	North Central province	0.02554977
β_5	Central province	0.11607572
β_6	Sabaragamuwa province	0.00473791
β_7	Eastern province	0.00972736
β_8	Uva province	0.00020758
β_9	Southern province	0.11356019

3.6. Impact of human dispersal behaviors into disease spread

The human dispersal behaviors were restricted by imposing lockdowns to decrease the COVID-19 spread in Sri Lanka. Hence a lockdown variable; L_i is introduced to demonstrate the impact of human dispersal behaviors on disease spread. The lockdown strength of i^{th} province is given by L_i which takes the values between 0 and 1. When the lockdown strength becomes 1, it denotes no human dispersal behaviors [43]. According to the literature, it causes a new transmission rate in province i $(\widetilde{\beta}_i)$ such that, $\widetilde{\beta}_i = (1 - L_i)\beta_i$.

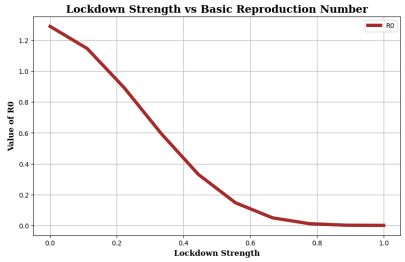

Figure 6: Basic reproduction number for different lockdown strengths.

4. Discussion and Conclusions

Epidemic modeling by a gravity model approach is getting more attention and it is becoming interesting to study among the researchers. When it comes to meta-population modeling using compartmental models, the gravity model used in Equation 3, could not be used to model the human dispersal behaviors within a patch. It only models the human dispersal behaviors between patches. Because the denominator of the gravity model Equation 3 becomes zero as explained in Section 2.1, when the human dispersal behaviors within a patch are considered. By modifying the gravity model as the Equation 6, the barriers to model the human dispersal behaviors within a patch disappear.

The human dispersal behaviors within a patch are observed to be greater than the behaviors between the patches in actual mobility data. Choosing the distance decay function $\mathrm{f}_\gamma(d_{ij})$ in the gravity model as $\dfrac{1}{(1+d_{ij})^\gamma}$ instead of $\dfrac{1}{(d_{ij})^\gamma}$ effectively captures the above-mentioned scenario and it also solves the problem of division by zero discussed in Section 2.1.

The distances between any provinces i and j which is denoted as d_{ij} is greater than 50km in Sri Lanka. ($d_{ij} > 50\text{km}, \forall i, j \in \{1, 2, ..., 9\}$ where $i \neq j$). Therefore,

$$\mathrm{f}_\gamma(d_{ij}) = \begin{cases} \dfrac{1}{(1+d_{ij})^\gamma} < \dfrac{1}{51^\gamma}, & \text{if } i = j \\ \dfrac{1}{(1+d_{ii})^\gamma} = 1, & \text{if } i \neq j \end{cases}$$

$\forall i \in \{1, 2, ..., 9\}$ where $\gamma = 0.77417147$.

Hence, $\mathrm{f}_\gamma(d_{ii}) > \mathrm{f}_\gamma(d_{ij}), \forall i, j \in \{1, 2, ..., 9\}$ where $i \neq j$. Therefore, the magnitude of the effect given by the distance decay function related to the human dispersal behaviors within a patch is greater than the magnitude of the effect given by the distance decay function related to the behaviors between patches. That explanation justifies the reason for capturing the higher human dispersal behavior rates within a patch by the modified gravity model in Equation 6. That scenario is represented visually in Figure 1.

The parameters of the modified gravity model in Equation 6 are estimated using three different methods as discussed in Section 3.1. Three different model evaluation metric values are calculated for all those sets of parameters obtained by above mentioned types of estimation methods. The R-squared, RMSE (Root Mean Square Error), and MAE (Mean Average Error) values are calculated for each model that carries different parameter values due to the difference in the estimation method.

Table 5: R-squared, RMSE, and MAE values across different estimation methods.

	OLS method	NLS method	Hybrid OLS-NLS method
R-squared	0.8858	0.7868	0.8963
RMSE	0.0799	0.1091	0.0761
MAE	0.0378	0.0546	0.0365

The lowest R-squared value and the highest RMSE and MAE values can observed in the model that uses the NLS method to estimate parameters. It occurs because of the high sensitivity to the initial guesses. At the same time, the hybrid OLS-NLS method discussed in sub-section 2.2.4, provides the highest value for R-squared and the lowest values for the RMSE and MAE. Therefore, parameter estimation using the hybrid OLS-NLS method gives better parameter values for the gravity model compared to the other two methods. Therefore, in this study, the parameters estimated using the hybrid OLS-NLS method are used as the parameters for the modified gravity model in Equation 6 which are $(0.38326244, 0.60072182, 0.77417147, 1.00976875)$ for (n, m, γ, K).

The parameter estimation method used in this research to estimate the gravity model parameters is the hybrid OLS-NLS method that combines the OLS and NLS methods. Since the parameters must be estimated using the OLS method with the log-linearised gravity model, only non-zero mobility included data can be used for the parameter estimation.

The relationship between the dependent variable F_{ij} and the independent variables P_i, P_j, and d_{ij} of the modified gravity model in Equation 6, can be displayed using the double-log functional form of the modified gravity model. The double-log functional form of it can be written as:

$$\ln(F_{ij}) = \ln K + n \ln(P_i) + m \ln(P_j) - \gamma \ln(1 + d_{ij}).$$

Then the n, m, and γ are the elasticity coefficients that represent how the dependent variable is affected by a change in independent variables.

It suggests that a 1% increase in the population of the residence province i causes a 0.38326% increase in the number of human movements from province i to province j. Also, it suggests that a result of a 0.60072% increase in the number of human movements from province i to j is caused by a 1% increase in the population of the province j. But a 1% increase in the distance decay function such that, $\mathrm{f}_\gamma(d_{ij}) = \dfrac{1}{(1+d_{ij})^\gamma}$ makes a 0.77417% decrease in the number of human movements from province i to j.

Then the mobility rates can be calculated using the equation introduced in Equation 10. Hence,

$$\pi_{ij} = \frac{K\dfrac{P_i^n P_j^m}{(1+d_{ij})^\gamma}}{\displaystyle\sum_{j=1}^{9} K\frac{P_i^n P_j^m}{(1+d_{ij})^\gamma}}, \qquad \forall i \in \{1,2,...,9\}.$$

Therefore, the proportionality constant K can be canceled off in the stage of the mobility rates calculation and it does not affect the mobility rates because,

$$\pi_{ij} = \frac{\dfrac{P_i^n P_j^m}{(1+d_{ij})^\gamma}}{\displaystyle\sum_{j=1}^{9} \frac{P_i^n P_j^m}{(1+d_{ij})^\gamma}}, \qquad \forall i \in \{1,2,...,9\}.$$

In Sri Lanka, there is a lack of mobility data. Only the data on buses are available as the mobility data. But there are many other traveling modes commonly used in Sri Lanka. Therefore, the parameter estimation for the gravity model is done with a limited number of mobility data.

The variation of the basic reproduction number with the lockdown strength can be observed in Figure 6. It shows that the reproduction number will be reduced when the lockdown strength of all provinces become increased which provides the decrease of the human dispersal behaviors. Therefore, the reduction of human dispersal behaviors causes to the reduction of the basic reproduction number. The disease die out when the human dispersal behavior become reduced to 80% because $R_0 < 1$ when the lockdown strength is 0.2.

In the initial stages of infectious disease, it is not practical to impose a full lockdown for a long time. Because it may cause various social and economic problems. Therefore the country is clustered into 3 clusters

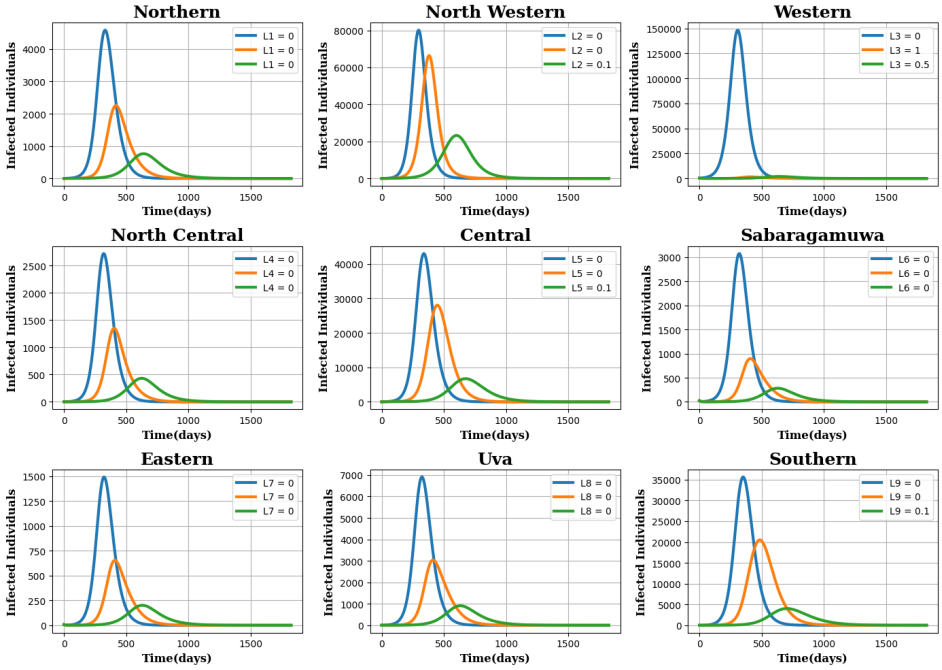

Figure 7: The infected human population over time in no lockdown, high zone lockdown, and high and moderate zone lockdown.

namely high-risk zone, moderate-risk zone and low-risk zone based on the number of infected humans in each province at their peak. The clustering is done using the K-Mean clustering.

Table 6: Clusters.

Zone	Provinces
High-risk	Western
Moderate-risk	Southern, North Western, Central
Low-risk	Other provinces

This clustering provides the opportunity to impose separate lockdowns with different strengths into separate clusters. Figure 7 shows infected curves for different lockdown scenarios for each province.

The infected curves when there is no lockdown are shown in blue colored curves and the orange color curves represent the infected curves when there is a full lockdown in the high-risk zone. The green-colored curves provide the infected curves when there are lockdowns with 0.5 and 0.1 lockdown strengths in high-risk zone and moderate-risk zone respectively. The least peaks for each province are provided by the last lockdown scenario.

5. Future Work

Since the gravity model gives an effective insight into modeling the patterns of human dispersal behaviors, this research can be extended to model the effects of the lockdown on the human dispersal behaviors and the patterns of human mobility amidst a lockdown incorporating the gravity model.

A transformation of the mobility matrix in Figure 1 into another mobility matrix which indicates the lockdown effects on the mobility within and between the provinces will be a new approach to model the lockdown effects on human dispersal behaviors. The transformation can be considered to be done using another matrix and multiplying it with the existing mobility matrix. Finding such a matrix that incorporates the lockdown measures will be an extension of this research.

In this research, a multi-patch SIR model with three compartments: susceptible, infected, and recovered are used. We can extend the multi-patch SIR model by incorporating another compartment or compartments such as dead, quarantined, and vaccinated compartments. It would be a more realistic capture of the real-world scenario of COVID-19 transmission. It will also open the research to a path of an extended version of epidemic modeling research incorporated with optimal control problems.

6. Data Availability

The population data for each province is taken from the Registrar General's Department of Sri Lanka [44]. The distances between the provinces are taken as mentioned in the Road Development Authority [45]. The mobility data is provided by the National Transport Commission of Sri Lanka. The COVID-19 situation reports from 15^{th} of November 2020 to 10^{th} of December 2020 are used as the data for number of COVID-19 patients [46].

References

[1] Centers for Disease Control and Prevention, Epidemiology Glossary-Data and Statistics-Reproductive Health, CDC, 2019. https://www.cdc.gov/reproductivehealth/data_stats/glossary.html, Accessed on April 12, 2024.

[2] Martcheva, M., An introduction to mathematical epidemiology, Springer, 61, pp. 9-31, 2015.

[3] Piret, J. and Boivin, G., Pandemics throughout history, Frontiers in Microbiology, 11, p. 631736, 2021.

[4] National Geographic Education, Epidemic. https://education.nationalgeographic.org/resource/epidemic/, Accessed on March 11, 2024.

[5] Chowell, G., Sattenspiel, L., Bansal, S. and Viboud, C., Mathematical models to characterize early epidemic growth: A review, Physics of Life Reviews, 18, pp. 66-97, 2016.

[6] Kretzschmar, M. and Wallinga, J., Mathematical models in infectious disease epidemiology, Modern Infectious Disease Epidemiology: Concepts, Methods, Mathematical Models, and Public Health, Springer, pp. 209-221, 2010.

[7] Basavarajaiah, D. and Murthy, B.N., COVID transmission modeling: An insight into infectious diseases mechanism, Chapman and Hall/CRC, 2022.

[8] Becker, N., The uses of epidemic models, Biometrics, pp. 295-305, 1979.

[9] Choisy, M., Guégan, J.F. and Rohani, P., Mathematical modeling of infectious diseases dynamics, Encyclopedia of Infectious Diseases: Modern Methodologies, 379, 2007.

[10] Zhang, M., Wang, S., Hu, T., Fu, X., Wang, X., Hu, Y., Halloran, B., Li, Z., Cui, Y., Liu, H. and Liu, Z., Human mobility and covid-19 transmission: a systematic review and future directions, Annals of GIS, 28(4), pp. 501-514, 2022.

[11] Balcan, D., Gon¸calves, B., Hu, H., Ramasco, J.J., Colizza, V. and Vespignani, A., Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model, Journal of Computational Science, 1(3), pp. 132-145, 2010.

[12] Werner, P.A., Kesik-Brodacka, M., Nowak, K., Olszewski, R., Kaleta, M. and Liebers, D.T., Modeling the spatial and temporal spread of covid-19 in poland based on a spatial interaction model, ISPRS International Journal of Geo-Information, 11(3), p. 195, 2022.

[13] Sen, A. and Smith, T.E., Gravity model of spatial interaction behavior in transportation networks, Transportation Research Record, 2302(1), pp. 130-139, 2012.

[14] Yang, X., Fang, Z., Xu, Y., Yin, L., Li, J., and Lu, S., Spatial heterogeneity in spatial interaction of human movements-Insights from large-scale mobile positioning data, Journal of Transport Geography, 78, pp. 29-40, 2019.

[15] Akuno, A.O., Ramírez-Ramírez, L.L., Mehta, C., Krishnanunni C., Bui-Thanh, T. and Montoya J.A., Multi-patch epidemic models with partial mobility, residency, and demography, Chaos, Solitons & Fractals, 173, p. 113690, 2023.

[16] Iyaniwura, S., Ringa, N., Adu, P.A., Mak S., Janjua, N.Z., Irvine, M.A. and Otterstatter, M., Understanding the impact of mobility on COVID-19 spread: A hybrid gravity-metapopulation model of COVID-19, PLOS Computational Biology, 19(5), p. e1011123, 2023.

[17] Arfken, G. International edition university physics, Elsevier, 2012.

[18] Greenwood, M.J., Modeling migration, Encyclopedia of Social Measurement, K. Kempf-Leonard (Ed.), Elsevier, pp. 725-734 2005.

[19] Frimpong, A., Arhin K., Boachie, M.K. and Acheampong, K., A gravity model approach to understand the spread of pandemics: Evidence from the COVID-19 outbreak, Open Health, 4(1), p. 20220032, 2023.

[20] Hong, I., Jung W.S. and Jo, H.H., Gravity model explained by the radiation model on a population landscape, PloS One, 14(6), p. e0218028, 2019.

[21] De Benedictis, L. and Taglioni, D., The gravity model in international trade, Springer, 2011.

[22] Anderson, J.E., The gravity model, Annu. Rev. Econ., 3(1), pp. 133-160, 2011.

[23] Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J.J. and Vespignani, A., Multiscale mobility networks and the spatial spreading of infectious diseases, Proceedings of the National Academy of Sciences, 106(51), pp. 21484-21489, 2009.

[24] Sallah, K., Giorgi, R., Bengtsson, L., Lu, X., Wetter, E., Adrien, P., Rebaudet, S., Piarroux, R., and Gaudart, J., Mathematical models for predicting human mobility in the context of infectious disease spread: introducing the impedance model, International Journal of Health Geographics, 16, pp. 1-11, 2017.

[25] Puangsun, S. and Patanarapeelert, K., An SIR epidemic model with gravity in patchy environment: Analyses for two patches system, KMITL Science and Technology Journal, 12(2), pp. 127-133, 2012.

[26] Viboud, C., Bjørnstad, O.N., Smith, D.L., Simonsen, L., Miller, M.A., and Grenfell, B.T., Synchrony, waves, and spatial hierarchies in the spread of influenza, Science, 312(5772), pp. 447-451, 2006.

[27] Li, X., Xu, S., Yu, M., Wang, K., Tao, Y., Zhou, Y., Shi, J., Zhou, M., Wu, B., Yang, Z. and Zhang, C., Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan, Journal of Allergy and Clinical Immunology, 146(1), pp. 110-118, 2020.

[28] Shi, Y., Wang, G., Cai, X.P., Deng, J.W., Zheng, L., Zhu, H.H., Zheng, M., Yang, B. and Chen, Z., An overview of COVID-19, Journal of Zhejiang University Science B, 21(5), pp. 343-360, 2020.

[29] World Health Organization, COVID-19 deaths-WHO COVID-19 dashboard, 2024. https://data.who.int/dashboards/covid19/deaths?n=c, Accessed on April 11, 2024.

[30] Amaratunga, D., Fernando, N., Haigh, R. and Jayasinghe, N., The COVID-19 outbreak in Sri Lanka: a synoptic analysis focusing on trends, impacts, risks and science-policy interaction processes, Progress in Disaster Science, 8, p. 100133, 2020.

[31] Epidemiology Unit, National epidemiological report - Sri Lanka, 2024. https://www.epid.gov.lk/high-endemic-diseases-outbreaks/high-endemic-diseases-outbreaks. Accessed on April 12, 2024.

[32] Adikari, K. and Dunusinghe, P., The Impact of Covid-19 Pandemic on Apparel Exports: Evidence from Sri Lanka, Indian Journal of Applied Economics and Business, 2023.

[33] Ganegoda, N.C., Perera, S.S.N., Götz, T., Wijaya, K.P., Jayathunga, Y., Peiris, H.O.W., Meththananda, R.G.U.I., Erandi, K.K.W.H., Gamachchige, R.N. and Thrimavithana, R.T., Mitigating covid-19 transmission in Sri Lanka: A mathematical outlook for decision support, 2021.

[34] Chen, Y., The distance-decay function of geographical gravity model: Power law or exponential law?, Chaos, Solitons & Fractals, 77, pp. 174-189, 2015.

[35] Miller, H.J., Tobler's first law and spatial analysis, Annals of the Association of American Geographers, 94(2), pp. 284-289, 2004.

[36] Arbia, G., Benedetti, R. and Espa, G., Effects of the MAUP on image classification, Geographical Systems, pp. 123-141, 1996.

[37] Cosner, C., Beier, J.C., Cantrell, R.S., Impoinvil, D., Kapitanski, L., Potts, M.D., Troyo, A., and Ruan, S., The effects of human movement on the persistence of vector-borne diseases, Journal of Theoretical Biology, 258(4), pp. 550-560, 2009.

[38] Brauer, F., Castillo-Chavez, C. and Feng, Z., Epidemiological models incorporating mobility, behavior, and time scales, Mathematical Models in Epidemiology, pp. 477-504, 2019.

[39] Department of Census and Statistics - Sri Lanka, Census map - Sri Lanka, 2023. http://map.statistics.gov.lk:8080/LankaStatMap/apps/gs/censusmap/index.html, Accessed on October 25, 2023.

[40] Diekmann, O., Heesterbeek, J.A.P., and Metz, J.A.J., On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, Journal of Mathematical Biology, 28, pp. 365-382, 1990.

[41] Heffernan, J.M., Smith, R.J., and Wahl, L.M., Perspectives on the basic reproductive ratio, Journal of the Royal Society Interface, 2(4), pp. 281-293, 2005.

[42] Erandi, K., Mahasinghe, A.C., Perera, S. and Jayasinghe S., Effectiveness of the strategies implemented in Sri Lanka for controlling the COVID-19 outbreak, Journal of Applied Mathematics, 2020, pp. 1-10, 2020.

[43] Lee, S., and Castillo-Chavez, C., The role of residence times in two-patch dengue transmission dynamics and optimal strategies, Journal of Theoretical Biology, 374, pp. 152-164, 2015.

[44] Registrar General's Department - Sri Lanka, Vital statistics, 2024. https://www.rgd.gov.lk/web/index.php?option=com_content&view=article&id=136&Itemid=294&lang=en#estimates-on-mid-year-poplulation, Accessed on April 12, 2024.

[45] Road Development Authority - Sri Lanka, EE Divisions, 2024. http://www.rda.gov.lk/source/distance_cities.htm, Accessed on April 12, 2024.

[46] Epidemiology Unit - Sri Lanka, Situation report, 2020. https://www.epid.gov.lk/situation-report?filter=2020, Accessed on April 12, 2024.