Spatial Assessment of COVID-19 First-Wave Mortality Risk in the Global South

Shawky Mansour, Ammar Abulibdeh, Mohammed Alahmadi & Elnazir Ramadan

To cite this article: Shawky Mansour, Ammar Abulibdeh, Mohammed Alahmadi & Elnazir Ramadan (2022): Spatial Assessment of COVID-19 First-Wave Mortality Risk in the Global South, The Professional Geographer, DOI: 10.1080/00330124.2021.2009888

To link to this article: https://doi.org/10.1080/00330124.2021.2009888
The coronavirus disease (COVID-19) that appeared in 2019 gave rise to a major global health crisis that is still topping global health, socioeconomic, and intervention program agendas. Although the outbreak of COVID-19 has had substantial and devastating impacts on developed countries, the countries of the Global South share a higher proportion of the epidemic’s effects as shown particularly in morbidity and mortality rates in low-income countries. Modeling the effects of underlying factors and disease mortality is essential to plan effective control strategies for disease transmission and risks. The relationship between COVID-19 mortality rates and sociodemographic and health determinants can highlight various epidemic fatality risks. In this research, geographic information systems (GIS) and a multilayer perceptron (MLP) artificial neural network (ANN) were adopted to model and examine variations in COVID-19 mortality rates in the Global South. The model’s performance was tested using statistical measures of mean square error (MSE), root mean square error (RMSE), mean bias error (MBE), and the coefficient of determination ($R^2$). The findings indicated that the most important variables in explaining spatial mortality rate variations were the size of the elderly (sixty-five and older) population, accessibility to handwashing facilities, and hospital beds per 1,000 population. Mapping the explanatory variables and estimated mortality rates and determining the importance of each variable in explaining the spatial variation of COVID-19 death rates across countries of the Global South can shed light on how public health care and demographic structures can offer policymakers invaluable guidelines to planning effective intervention strategies. 

Key Words: ANN, COVID-19 mortality, GIS modeling, Global South, sociodemographic determinants.

A series of pneumonia cases with unidentified etiology emerged in Wuhan (Hubei), China. COVID-19 is a new category of coronavirus that has a very dynamic structure and spreads rapidly by transmission from human to human. In the past decade, two zoonotic coronaviruses spread globally: severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV; Breban, Riou, and Fontanet 2013; Chan et al. 2015; Vijay and Perlman 2016; Zhou et al. 2016). More than 322 million people have been infected so far by COVID-19 in more than 200 countries, resulting in more than 4.7 million deaths as of September 2021. These numbers are increasing every day and the number of infected cases worldwide is also constantly increasing. The virus is currently evolving and spreading in many middle- and low-income countries after spreading in China, Europe, and North America (Bong et al. 2020; Roberton et al. 2020; Thiemann et al. 2020). A collaborative and coordinated global response is therefore needed to prepare health systems to overcome this unprecedented challenge.

Prior to development of a vaccine, the best practice to avoid infection and control disease transmission is through prevention and management procedures (Abulibdeh 2020; Adhikari et al. 2020). This requires collective efforts by governments and public. There are a number of ways to avoid exposure to the virus, including avoiding contact with infected people and social distancing (Dalton, Corbett, and Katelaris 2020; World Health Organization [WHO] 2020), and some actions taken globally to control the spread of the disease have risked socioeconomic disorder and confusion. They include closures of schools, universities, businesses,
and public institutions; banning travel; self-quarantine; social isolation; national lockdowns; and cancellation of sports activities (Brzezinski et al. 2020; Lau et al. 2020; Abulibdeh 2021).

To control the COVID-19 pandemic and to protect the global community from the threat of the virus, many countries accelerated developing, manufacturing, and distributing an effective vaccine. Different types of vaccines were developed (Pfizer–BioNTech, Moderna, Oxford–AstraZeneca, and Janssen [Johnson & Johnson]) and approved by many countries across the globe. To control the spread of the disease, there is a need to immunize an extraordinarily large number of people in each country. Global vaccine distribution remains highly unequal, however, and most of the supplies have been distributed in high-income countries in North America and Europe (Solís Arce et al. 2021; Wouters et al. 2021).

The United Nations (UN) identified “Ensure healthy lives and promote well-being for all at all ages” as the third of its Sustainable Development Goals (SDGs) to be reached by 2030, with the aim of improving the health of millions of people across the globe (WHO 2015; Stenberg et al. 2017). Nevertheless, the global response to this pandemic has not been organized as a collective effort, with governments’ response to the pandemic being based on their capabilities, expertise, possibilities, simulations, and spatial modeling (Hale et al. 2020; Tanne et al. 2020). Variations in capacity to test and hospitalize; in numbers of physicians, nurses, and midwives; and in health infrastructure make it difficult for many countries to identify accurately the number of infected people. These differences are found even among industrialized countries. Kontis et al. (2020) compared the effect of the pandemic on mortality rates in nineteen industrialized countries in Europe, Australia, and New Zealand using sixteen Bayesian models. They found that the pandemic increased the mortality rate in these countries; however, there was a high variation in the mortality rate between these countries. These variations are due to the population and community characteristics, resilience and preparedness of the public health system, immediate response to the disease, and health and social care systems in each country.

In the Global North, trillions of dollars have been committed to strengthening health systems and overcoming the negative impact on social and economic sectors, whereas limitations in health systems and especially in health facilities coverage mean that most Global South countries suffer socially and economically from COVID-19 and show high reported case numbers and mortality rates (Cash and Patel 2020; Shet et al. 2020). Moreover, many countries in the Global South are unprepared to manage an outbreak of the disease and, consequently, high prevalence rates, infection risks, and mortalities are expected. Most low-income countries are in fragile and conflict-affected situations and experiencing social, economic, and political problems and crises that COVID-19 can only exacerbate. In addition, many of these countries, particularly in Africa, already experience debt distress, and interest payments restrict the resources needed to improve social and health protection (Bong et al. 2020; Dahab et al. 2020; Mehta et al. 2020). For instance, several countries in Africa are spending more than three times what they spend on their health systems to service external debt (Action Aid International 2020). This calls into question whether, in many of these countries, the national health system has the capacity to respond effectively to critical infected cases requiring intensive care for COVID-19 pneumonia. Furthermore, many Global South countries lack comprehensive social security systems and hence have some difficulties in developing a robust financial rescue package for the health system. These countries are therefore more vulnerable to exogenous shocks and new waves of the current pandemic and are unprepared and unable to afford the viral emergency, with catastrophic consequences for their health systems (Action Aid International 2020).

Different methods and models have been used to predict the spread of COVID-19 prevalence, incidence, and mortality rates in different countries and the associated factors that increase the propagation of the pandemic (Dudel et al. 2020; Verity et al. 2020). Li, Feng, and Quan (2020) predicted the ongoing trend and the outbreak size of COVID-19 in China by developing a function using data-driven analysis. Fanelli and Piazza (2020) used simple mean-field models and built iterative time lag maps to analyze the temporal dynamics of COVID-19 disease outbreak in China, Italy, and France. Roda et al. (2020) used the Akaike information criterion to compare SIR and SEIR (susceptible, exposed, infective, recovered) models in modeling COVID-19 spread in Wuhan Province, China, considering the quarantine measures undertaken in the city. Al-Qaness et al. (2020) used a modified adaptive neuro-fuzzy inference system to predict the number of infected cases of COVID-19 in ten days based on previously confirmed cases recorded in China. Ceylan (2020) used autoregressive integrated moving average to predict the epidemiological trend of COVID-19 prevalence in Italy, Spain, and France.

Spatial distribution models are crucial to statistically examine the geographic correlation and independences between several explanatory variables and disease outbreaks such as a COVID-19 pandemic (Mollalo, Vahedi, and Rivera 2020). Various spatial modeling techniques within a geographical information systems (GIS) environment have been adopted to investigate, analyze, visualize, and map the COVID-19 pandemic’s spatial distribution (e.g., Kamel Boulos and Geraghty 2020; Mollalo, Vahedi, and Rivera 2020; Sarwar et al. 2020). For example, GIS methods were used to model the spatial distribution of infectious diseases (Lovett et al. 2014;
Mollalo, Vahedi et al. 2018; Mollalo, Mao et al. 2019), which can help identify the most effective places in which to combat the disease’s spread. Some studies have used GIS-based analysis to investigate the spatial distribution of COVID-19 since the outbreak of the disease. Conducting spatial local modeling, Mollalo, Vahedi, and Rivera (2020) used environmental, topographic, demographic, and socioeconomic explanatory variables to examine the spatial distribution of COVID-19 incidence across the United States at the county level. They also examined the spatial dependency using spatial error models as well as investigating local spatial nonstationarity using geographically weighted regression and multiscale geographically weighted regression models. Kamel Boulos and Geraghty (2020) presented practical online and mobile GIS and mapping dashboards and applications such as Johns Hopkins University’s The Center for Systems Science and Engineering, the WHO dashboard, HealthMap, WorldPop, and EpiRisk to track the COVID-19 pandemic around the world. Arif and Sengupta (2021) found a positive spatial correlation between COVID-19 transmission and population density in four south Indian states using the Thiessen polygon in a GIS environment.

Few studies have investigated the factors that contribute to the spread of the COVID-19 pandemic in the Global South. For example, Chowdhury and Jomo (2020) reviewed the strategies adopted by some countries in the Global South to reduce the spread of the disease and the lessons drawn from these policies. Al-Ali (2020) investigated the gendered implications of COVID-19 in the Global South considering preexisting gender inequalities. Oldekop et al. (2020) argued that there is a need for a global development paradigm and the need for a more sustainable and equitable world as a result of the spread of the pandemic. Porter et al. (2021) investigated the impact of the pandemic on the mental health of young people in some countries in the Global South. Shamasunder et al. (2020) discussed the inequality in the global health system and analyzed the effect of these inequalities in the context of the pandemic. None of the studies that investigated the spread of the COVID-19 pandemic focused on the spatial distribution of COVID-19 mortality rates considering demographic and health explanatory variables. Therefore, this study fills the gap by investigating the impact of demographic and health factors on the mortality rate in the Global South.

Modeling the effects of sociodemographic factors and health system facilities on the COVID-19 pandemic in the Global South countries has not been carried out yet. Therefore, the overarching aim of this study is to model the effects of several sociodemographic and health drivers on COVID-19 mortality across the Global South countries using a GIS-based artificial neural network (ANN) technique. To understand patterns of disease transmission and fatalities, the study also explicitly estimates the global spatial distribution of COVID-19 mortality. To the best of our knowledge, this research provides the first attempt to address global geographic modeling of COVID-19 in the Global South. This analysis supports decision makers and can provide useful insights for health authorities and organizations as well as for policymakers to target interventions and allocate adequate resources including intensive care facilities, public health equipment, and hospital beds to mitigate disease transmission and mortality risks in the Global South and particularly in low-income countries.

Data and Materials

Global South is a term describing countries located in Asia, Africa, Latin America, and the Caribbean and classified by the World Bank as low- or middle-income countries. This term is the most recent in a long line of concepts used to define, identify, and cluster low- and middle-income countries or the “poorer parts of the world” (Mahler 2017). Other terms used to describe these countries include third world, developing, less developed, and underdeveloped countries (Muni 1979). Some scholars acknowledge that geography and geopolitical relations are important but still also acknowledge a growing gap in wealth and power within countries (Naudé 2010). Therefore, this term is used to express the uneven national social and economic development as well as political differences, positions, and experiences between the countries in this world.

Demographic and health explanatory variables were used in this study to model the spatial distribution of COVID-19 mortality rates in the Global South. The explanatory variables as well as COVID-19 mortality rate were acquired from secondary data sources, particularly the UN, Division of Statistics (2020) and Worldometers (2021) and combined with the spatial layer to create the whole geodatabase of modeling process. Three demographic variables were generated to represent the geographic distribution of elderly population (age sixty-five and older), population density, and accessibility to basic handwashing facilities including soap and water (Table 1). The purpose of using these three variables is to examine the potential effects of demographic factors on COVID-19 mortality rates and whether high mortality is highly concentrated in this age cohort. Population density is associated with disease prevalence and transmission in Global South countries; the higher the population density, the higher the disease transmission rate in a population living in an overcrowded environment. The measure of handwashing with soap and water is an important hygiene behavior...
<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Definition</th>
<th>Connection to effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital beds (per 1,000 people)</td>
<td>Hospital beds include inpatient beds available in public, private, general, and specialized hospitals and rehabilitation centers</td>
<td>Hospital beds is an indicator of the availability of inpatient services and the strength of a country’s health care system. The more hospital beds, the lower the mortality rate (inverse relationship; Berman 2000; Delamater et al. 2013).</td>
</tr>
<tr>
<td>Current health expenditure (% of GDP)</td>
<td>The total public health care expenditure as percentage of GDP</td>
<td>Allocating and mobilizing funds to strengthen the health system with the required resources might mitigate the spread of the pandemic as well as the associated negative impact on well-being. The greater the health expenditure, the lower the mortality rate (inverse relationship; Bhalotra 2007; Siroka, Ponce, and Lonnroth 2016).</td>
</tr>
<tr>
<td>People with basic handwashing facilities including soap and water (% of population)</td>
<td>The percentage of the population living in households that have a handwashing facility with soap and water</td>
<td>Hand hygiene is one method of preventing infection by the virus that causes COVID-19. The more handwashing practices, the fewer infected COVID-19 cases and the lower mortality rate (inverse relationship; Adhikari et al. 2020).</td>
</tr>
<tr>
<td>Tests per 1 million population</td>
<td>The percentage of tests that are performed to identify COVID-19 cases per 1 million people</td>
<td>The number of tests is significant in interpreting the data on confirmed cases and allowing the authorities to respond appropriately to monitor and mitigate the transmission of the disease (Ghosal et al. 2020 The more tests, the more discovered COVID-19 infected cases but the lower the mortality rate (inverse relationship).</td>
</tr>
<tr>
<td>Incidence of TB (per 100,000 people)</td>
<td>The estimated number of new and relapse cases arising in a given year, expressed as the rate per 100,000 population</td>
<td>TB affects the lungs and hence infection with COVID-19 might worsen the condition of the TB patient (Chen et al. 2020). Patients with TB are more vulnerable to COVID-19 (direct relationship).</td>
</tr>
<tr>
<td>Nurses and midwives (per 1,000 people)</td>
<td>The percentage of nurses and midwives per 1,000 people</td>
<td>Nurses and midwives are an important indicator of the health system’s capacity and its ability to deal with the spread of the disease (Eckert 2020). The greater the number of nurses and midwives, the lower the mortality rate (inverse relationship).</td>
</tr>
<tr>
<td>Infected critical cases</td>
<td>The number of infected COVID-19 critical cases that need hospitalization and ICU services</td>
<td>The number of critical cases could increase the burden on the health care system, particularly on ICU and ventilation systems. The more critical cases, the higher the mortality rate (direct relationship).</td>
</tr>
<tr>
<td>Physicians (per 1,000 people)</td>
<td>The percentage of physicians per 1,000 people</td>
<td>Physicians are significant for creating a robust health care system and can play a major role in mitigating the spread of the COVID-19 pandemic (Tanne et al. 2020). The more physicians in the health care system, the lower the mortality rate (inverse relationship).</td>
</tr>
<tr>
<td>Recovered cases</td>
<td>The number of infected COVID-19 cases that have recovered from the disease</td>
<td>The number of recovered cases compared to the number of mortalities in a country could be an indicator of the robustness of the health care system and the effectiveness of mitigation measures taken by the authorities. The more recovered cases, the lower the mortality rate (inverse relationship).</td>
</tr>
<tr>
<td>Incidence of HIV (per 1,000 uninfected population age 15–49)</td>
<td>Number of new HIV infections among uninfected populations aged 15–49 expressed per 1,000 uninfected population in the year before the period</td>
<td>HIV affects the immune system and results in gradual and persistent failure and decline of health so that people with HIV are more vulnerable to COVID-19 infection (Adepoju 2020). HIV disease increases the mortality rate in COVID-19 patients (direct relationship).</td>
</tr>
</tbody>
</table>

(Continued)
for human health and considered one of the top priorities and most cost-effective interventions to prevent the spread of the COVID-19 pandemic.

In addition to the demographic variables, health explanatory variables have an enormous impact on the magnitude of the spread of the COVID-19 pandemic and the mortality rate in Global South countries. In this study, nine health explanatory variables were considered, as shown in Table 1. These variables represent the health infrastructure (i.e., hospital beds; number of physicians, nurses, and midwives; and health expenditure), health preparedness (i.e., tests, critical cases, and recovered cases), and dominant disease (e.g., tuberculosis [TB] and human immunodeficiency virus [HIV]) that elevate critical cases. These variables are indicators of a country’s capability and preparedness to mitigate the spread of the disease and reduce its mortality rate. Furthermore, these variables are important in determining a country’s capability to strengthen the health care system with the required resources and the appropriate health system capacity ratio (Gilbert et al. 2020; Mansour et al. 2021). Identifying the geographic distribution of other preexisting comorbidities (e.g., TB and HIV) also might help in assessing the potential mortality rate for the COVID-19 pandemic and hence its expected burden on the health care system. Investigating the geographic distribution of these variables is therefore important in explaining the geographic distribution of the COVID-19 mortality rate in Global South countries.

During the first wave of the pandemic, the novel coronavirus spread rapidly throughout the world, causing hundreds of thousands of mortality cases. Figure 1 shows the variation in infected, recovered, and mortality cases between the continents and countries. For example, in Latin America, the highest number of infected cases and mortality rate are located in Brazil. Additionally, many countries in Latin America including Argentina, Bolivia, Peru, and Colombia have approximately the same mortality rate but different numbers of infected cases. In Africa, although the number of infected cases is low, a high mortality rate in many countries is exhibited. Most of these countries are located in the central or northern parts of the continent, such as Algeria, Chad, Niger, Mali, and Mauritania. In Asia, the number of infected cases and mortality rates vary between regions and countries. In East Asia, China has the highest number of infected cases and the highest mortality rate. In South Asia, the highest numbers of infected cases are in India and Pakistan, but the highest mortality rate is in India. In Southeast Asia, Indonesia and the Philippines have the most infected cases and the highest mortality rates. Finally, in the Middle East, high rates of infected cases are distributed in Iran and Saudi Arabia and the highest mortality rates are in Iran and Yemen. It is quite clear that the variations in geographic distribution of infected cases and mortality rates in the Global South are interesting, providing an incentive to investigate the causes of this variation among these countries.

### Methods

**Artificial Neural Network**

An ANN is a type of artificial intelligence and machine deep learning technique that mimics the human brain and particularly the neural system (Schmidhuber 2015; Balasaravanan and Prakash 2017; Mishra and Gupta 2017). The ANN architecture comprises a high number of perceptrons (neurons) that are used for classification and prediction processes when the relationships between the dependent and independent variables are unknown or nonlinear (G. Panchal et al. 2011; Ettayoub, Lazaar, and Ghanou 2013; F. S. Panchal and Panchal 2014).

Fundamentally, an ANN is characterized by three major components: architecture, learning algorithm, and activation function (Idri, Mbarki, and...
Abrahm 2004; Linares-Rodriguez et al. 2013). The most commonly used ANN techniques are multi-layer perceptron (MLP) and radial basis function (Bayram et al. 2015; Gholami et al. 2020). These techniques are used widely in various scientific applications and domains to address entangled relationships among environmental, socioeconomic, and demographic phenomena. For instance, ANN algorithms were used in disease medical decision support systems and disease diagnosis (e.g., Shanthi, Sahoo, and Saravanan 2008; Nayeem, Wan, and Hasan 2015; Hasan, Jasim, and Hashim 2017). Similarly, MLP techniques have been employed in examining the mortality risk of several diseases (Sibanda and Pretorius 2011, 2012; Süt and Çelik 2012; Moridani et al. 2015). A recent study has simulated mortality risks in patients with COVID-19 (Pourhomayoun and Shakibi 2020) where a global data set of 117,000 patients with laboratory-confirmed COVID-19 was analyzed using machine learning algorithms. The study focused only on patient symptoms and disease features, however, although spatial variations of causative predictors were included in the simulation process.

In this study, a predictive model has been developed spatially to estimate COVID-19 mortality incidence in Global South countries adopting twelve demographic and socioeconomic explanatory variables. The procedure adopted for the development of the ANN model consists of several stages.

**Developing an MLP Model.** In addition to the dependent variable that represents the percentage of mortality in each country, several demographic, socioeconomic, and health variables were compiled in a spatial aggregated database based on a national scale for training (learning) and testing of various ANN models. Spatial layers, attribute linkage, and mapping were conducted within a GIS environment. An MLP model was utilized where the network structure includes several inputs, hidden layers, and output layers. Each hidden layer receives the values from a specific number of input layers and then a weight is calculated before passing it to the output layer. Through the weighting process, the connections allow the entered variables to transfer through the layers, where each neuron receives data from the previous layer and calculates a weighted sum of all its net inputs as follows:

\[
y_i = f \sum_{j=1}^{n} w_j y_j + b_i, \tag{1}
\]

where \(y_i\) reveals the output at node \(i\) and \(f\) is an activation function; \(x\) is the input vector of inputs \((x_1, x_p, x_n)\); \(w\) is the weight vector \((w_1, \ldots, w_p, \ldots, w_n)\); and \(w_j\) is the weighted connection from the \(i\)th node in the preceding layer to node \(j\). \(b_i\) is a bias or
a constant that adjusts the output along with synaptic weights to optimize the model fit.

In an ANN, a neuron $N$ can be mathematically expressed as follows:

$$ u_N = \sum_{j=1}^{k} w_{kj} x_j, $$

where $u_N$ is the linear combination; $x_1, x_2, \ldots, x_n$ signify signals and input variables; and $w_{k1}, w_{k2}, \ldots, w_{kn}$ designate the synaptic weights of neuron $N$. The output of the neuron $N$ also can be expressed as follows:

$$ y_N = \theta(u_N - \theta_N), $$

where $\theta$ is the action function (linear or nonlinear) and $\theta_N$ is the threshold.

**ANN Transfer Function.** For this ANN model, a sigmoid activation function was adopted as the most commonly used transfer function. It is computed as follows:

$$ f(n) = \frac{1}{1 + e^{-x}} \quad 0 =< f(n) > 1, $$

where $x$ indicates the input value and $e$ specifies a constant (2.718, ...). In the sigmoid function the value of $e - x$ is always positive. Consequently, $1 + e - x$ always results in a value greater than 1 and 1 divided by $1 + e - x'$ (which is greater than 1) always ends up between 0 and 1.

**Neural Network Architecture.** In this research, an MLP was implemented within a GIS framework to estimated COVID-19 mortality rates across the Global South. The structure of three interconnected ANN layers is displayed in Figure 2. The network involves one input layer, two hidden layers, and one output layer. The input layer includes twelve nodes, each representing one explanatory variable, and the output layer represents the continuous estimated variable of COVID-19 mortality. Each node in the hidden layer receives synaptic weights after components of a special input vector are propagated by the input layer. Through a nonlinear-linear function (e.g., sigmoid or tan-sigmoid), the output is computed incorporating the weighted sums. In the neural network processing, we divided the data into

**Figure 2** Structure of artificial neural network designed for COVID-19 mortality estimation.
subsets, namely, training and testing data sets. About 75 percent of cases were incorporated as training data and 25 percent of the data were used as testing and validation data for the prediction model.

Accuracy of the ANN Model
In this step, a testing process was employed to assess the performance of the ANN model in estimating COVID-19 mortality rates. The root mean square error (RMSE) and coefficient of determination ($R^2$) were used as error statistics and goodness-of-fit measures to evaluate the models’ performance. These error terms and metrics were calculated using the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_i - p_i)^2},$$

where $a_i$ represents the actual value, $p_i$ denotes simulated value, and $n$ indicates the total number of observations. The RMSE is the square root of the average of sum of squared errors and it computes the average magnitude of the error between model predictions ($p$) and observed values ($a$). The coefficient of determination, $R^2$, was used to measure the predictive ability of the model and is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (d_i - p_i)^2}{\sum_{i=1}^{n} (d_i - \bar{a}_i)^2},$$

where $n$ illustrates the total number of observations and $a_i$ refers to the actual values. $P_i$ represents the predicted values and $\bar{a}_i$ signifies the average of the observed values.

Results
Model Test and Validation
The network structure was based on two hidden layers and different numbers of neurons that varied from five to thirty-five. The performance of various ANN models was compared and the optimal MLP model with its number of neurons was initially identified from seven configurations according to the minimum errors in three measures (mean square error [MSE], mean bias error [MBE], and RMSE) and the coefficient of determination, $R^2$ (Table 2). The first model, which has five neurons in the first and second hidden layers, shows the largest $R^2$ value (0.61), lowest MBE (0.01), and lowest RMSE (0.27).

To examine the quality of COVID-19 mortality prediction from the optimal ANN, the actual and predicted values of mortality percentages were plotted (Figure 3). The concentration of points around the line clearly indicates satisfactory performance of the selected model.

### Table 2: Results of different examined artificial neural network models with two hidden layers

<table>
<thead>
<tr>
<th>Neurons</th>
<th>MSE</th>
<th>RMSE</th>
<th>MBE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP 5–5</td>
<td>0.07</td>
<td>0.27</td>
<td>0.01</td>
<td>0.61</td>
</tr>
<tr>
<td>MLP 10–5</td>
<td>0.08</td>
<td>0.28</td>
<td>0.02</td>
<td>0.56</td>
</tr>
<tr>
<td>MLP 15–10</td>
<td>0.13</td>
<td>0.36</td>
<td>0.04</td>
<td>0.55</td>
</tr>
<tr>
<td>MLP 20–15</td>
<td>0.16</td>
<td>0.40</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>MLP 25–20</td>
<td>0.07</td>
<td>0.26</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>MLP 30–25</td>
<td>0.14</td>
<td>0.38</td>
<td>0.03</td>
<td>0.41</td>
</tr>
<tr>
<td>MLP 35–30</td>
<td>0.89</td>
<td>0.94</td>
<td>−0.10</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Note: MLP = multiplayer perceptron; MSE = mean square error; RMSE = root mean square error; MBE = mean bias error.

The Effects of Demographic Structure on COVID-19 Mortality
The results of the ANN model illustrate that the most important factor (Table 3) explaining the mortality rate is being sixty-five years of age or older, with ANN importance of 0.142. Figure 4C shows the geographic distribution of people aged sixty-five and above across the Global South. Higher percentages of this age cohort are concentrated mainly in Latin America and some Asian countries such as China, India, Iran, Thailand, and Malaysia, whereas this variable is less important in Africa except for South Africa and some countries in northern Africa, such as Morocco, Tunisia, and Egypt. This variable is an influential factor in explaining the mortality rate from COVID-19 in Brazil, Paraguay, Guyana, Colombia, and Ecuador in Latin America. The variable is also important in explaining the mortality rate in some Asian countries such as China, India, Iraq, Syria, Iran, Yemen, Thailand, and Malaysia, whereas the percentage of the population that is elderly is smaller in Africa except for Egypt, Algeria, and Tunisia.

The WHO identified handwashing as one method of preventing infection. This process is critical in preventing the transmission and spread of COVID-19. The ANN model shows that handwashing is the second most important variable in explaining the mortality rate in the Global South countries, with a coefficient of 0.130 as shown in Table 3. Figure 5B shows the geographic distribution of the percentage of households who have access to handwashing facilities in the Global South countries. Figure 5B shows that some countries in Latin American such as Brazil, Argentina, Peru, and Venezuela demonstrate low access to water and soap handwashing facilities. This trend is also found in many countries in Asia such as China, Iran, and some countries in the Middle East. In Africa, many countries in the central part of the continent in addition to Libya also show a low percentage of handwashing accessibility. This variable can explain the high mortality rate in these countries, particularly China and Brazil, as shown in Figure 1.
Population density is an important factor to be considered in analyzing the geographic distribution of the spread of COVID-19. This virus is spread by transmission from human to human; therefore, high population density could cause high rates of disease infection. Among the explanatory variables, population density is the eighth most important variable (coefficient = 0.064), as shown in Table 3. Population density is high across Asian countries, particularly in the eastern and southeastern regions of the continent (Figure 4D). The geographic distribution of population density in Africa shows high densities in some countries in the northern part, particularly Egypt, Morocco, and Tunisia; in the west, such as Nigeria and Ghana; and in the east, such as Ethiopia, Kenya, and Uganda. Hence, this variable is important in explaining the estimated mortality rates in many countries in Asia and Africa, whereas it is less important in Latin America.

### The Effects of Health Systems on COVID-19 Mortality

According to the MLP output, the effects of nine health variables on the COVID-19 mortality rate vary across the Global South countries. Hospital beds (per 1,000 people) is the most important health variable in explaining the mortality rate (coefficient = 0.095), as shown in Table 3. The geographic distribution of hospital beds in Global South countries (Figure 5A) shows a variation between these countries and between the Global South continents. Most countries in Latin America, particularly Argentina, have a high percentage of hospital beds. In Asia, China, Mongolia, Turkmenistan, and some countries in Southeast Asia and the Middle East (Gulf countries) have a high percentage of hospital beds per 1,000 population, whereas Iran has a smaller percentage. In contrast, a considerable number of countries in Africa show low numbers of
hospital beds as a percentage of the population. The significance of this factor is that a greater number of hospital beds gives the health care system greater capacity to treat critical cases and hence mitigate the mortality rate. Infected critical cases are the second most important health care variable in explaining the mortality rate (coefficient $\beta = 0.091$). The geographic distribution of the infected critical cases shows that many of these cases are concentrated in Latin American countries such as Brazil, Argentina, Chile, Peru, Colombia, and Ecuador. On the other hand, no information on infected critical cases is available for many countries in Africa, particularly those in the central part of the continent. Countries such as Tanzania, Madagascar, and Kenya in the east; Senegal and Guinea in the west; and South Africa in the south, however, have the highest number of critical cases on the continent (Figure 4A). In Asia, no information is available on critical cases in China. Critical cases are highest in India, Iran, Mongolia, and Thailand in the east and lower in countries in the Middle East. This is reflected in the estimated mortality rate, because countries with high numbers of critical cases such as Brazil and India have higher mortality rates.

Tuberculosis is a chronic disease that affects the lungs; hence, TB patients are more vulnerable to COVID-19 infections and the risk of mortality. Incidence of TB per 100,000 people is fifth among the variables used in this study to explain the COVID-19 mortality rate (coefficient $\beta = 0.091$). The geographic distribution of TB in Global South countries shows that it is spread widely in low-income countries in both Africa and Asia (Figure 5E). The prevalence of TB in the central and southern parts of Africa is higher, whereas in Asia higher percentages are found mainly in India, Pakistan, Afghanistan, Mongolia, and some countries in
Southeast Asia. In Latin American countries, the incidence of TB is very low. The incidence of TB might raise the mortality rate in many countries in Asia, such as India, Pakistan, Cambodia, Vietnam, Malaysia, the Philippines, and Indonesia, and in Africa, such as Tanzania, Zambia, Zimbabwe, Botswana, Namibia, and Angola, as shown in Figure 6. Another important variable is the incidence of HIV, which influences the immune system and causes its gradual failure. This disease is mainly concentrated in the central and southern parts of Africa (Figure 5F). The geographic distribution of this disease is very low in Latin America and in Asia. A higher percentage of this disease is found in Africa and it is correlated with COVID-19 mortality rate in the continent’s central and southern countries.

The increased rates of TB and HIV together are associated with high COVID-19 mortality rates (Figure 7A). This implies that the spread of COVID-19 in these parts of the continent will seriously threaten the lives of people there. Similarly, high health expenditure is associated with high recovery rates and low rates of mortality (Figure 7B). For example, Asian countries are characterized by large populations but have conducted a low number of COVID-19 tests. The number of tests is vital in reporting infected cases and allowing the authorities to respond appropriately to monitor and mitigate the transmission of the disease. Based on the ANN model, this variable is the sixth most important variable in explaining the mortality rate in Global South countries (coefficient

Figure 5 Distribution of six explanatory variables of COVID-19 mortality rates across the Global South. GDP = gross domestic product; HIV = human immunodeficiency virus.
The geographic distribution of COVID-19 tests in most low-income countries (Figure 5D) reveals a low number of conducted tests to identify infected cases. In Latin America, countries such as Chile, Peru, and Venezuela performed the highest number of tests per million population and the estimated mortality rates in these countries are low to moderate (Figure 6). In Africa, the number of tests per million population is very low except for a few countries such as Morocco, Tunisia, and South Africa. This might be attributed to insufficient testing capacity and low efficiency in detection. In Asia, the countries in the east and southeast have low numbers of tests, and some countries such as India and Pakistan show high projected mortality rates. The number of tests is also associated with the percentage of physicians, nurses, and midwives. The ANN importance coefficient of nurses and midwives variable is 0.078, whereas the coefficient of physicians is 0.056. The geographic distribution of

Figure 6 The modeled mortality percentages of COVID-19 in Global South countries.

Figure 7 Three-dimensional plot of COVID-19 mortality rate and (A) the incidence of tuberculosis and HIV and (B) health expenditure and recovery rate. HIV = human immunodeficiency virus.
physicians is extensive in Latin America and Asia, whereas in Africa higher percentages are concentrated in the northern part and particularly in Libya, Algeria, and Tunisia (Figure 4E). On the other hand, the geographic distribution of nurses and midwives shows a high rate of nurses and midwives in the Gulf Cooperation Council, followed by countries in East and Southeast Asia. In Africa, Tanzania, South Africa, Libya, Botswana, Uganda, and Senegal have the highest percentages of nurses and midwives. Likewise, Brazil in Latin America has a high number of nurses and midwives, followed by countries such as Argentina, Peru, Colombia, and Ecuador. Despite the high number of physicians, nurses, and midwives in most of the Global South countries, the mortality rate is still high in many of these countries. This explains the low importance of these two variables compared to the other health variables.

Government expenditure on public health care systems varies among Global South countries. According to the model, this variable is the ninth (Table 3) most important (coefficient = 0.06) in explaining the mortality rate in Global South countries. This is due to low spending of many countries across the three continents with less than 6 percent of their gross domestic product (GDP; Figure 5C). In Latin America, some countries (e.g., Brazil) spend more than 8.6 percent of their GDP on their health care systems. In Africa, several countries in the northern part (i.e., Algeria, Tunisia, Sudan, Libya, Niger) and southern part (i.e., South Africa, Namibia, Zimbabwe) allocate more than 8.6 percent of their GDP to the health care system, whereas most of the central and western African countries spend less on public health care. In Asia, Iran and the Gulf Cooperation Council countries, Turkmenistan, Afghanistan, China, and Mongolia spend more than 6 percent of their GDP on health infrastructure. Despite this spending on the health system as a percentage of GDP, however, the mortality rate is still high in many of these countries, particularly Iran. In the Middle East there is a recognized high recovery rate (Figure 4B), whereas in Latin America the recovery rate is lower than in Asia and some countries in Africa. This is due to the COVID-19 pandemic’s earlier incidence in Asia and Africa compared to Latin America. Examining the relationship between the mortality rate, recovery rate, and health expenditure reveals that the mortality rate will be high if the health expenditure is high but the recovery rate is low (Figure 7B) and that the mortality rate will be low if the health expenditure and recovery rate are both high.

**Discussion**

In this research, we examined the assumption that in low- and middle-income countries, health infrastructure and sociodemographic factors are significantly associated with mortality rates, particularly from communicable diseases such as COVID-19. Specifically, twelve explanatory variables were accumulated and an MLP ANN model was constructed to estimate the mortality rates across the Global South countries. Demographic characteristics can help in understanding how the COVID-19 pandemic has spread geographically in the Global South countries and why the mortality rate varies between these countries. Although all ages are susceptible to COVID-19 virus infection, the risk of mortality associated with this disease is higher among elderly people, mainly over sixty-five years of age. Older people are more vulnerable and need respiratory support as part of their treatment. The findings showed that, as expected, the proportion of the population aged sixty-five and older is a significant indicator of COVID-19 fatalities. The importance of age structure in explaining disease prevalence and risk mortality is consistent with previous analyses (e.g., Dudel et al. 2020; Verity et al. 2020). Different studies found that the mortality rate is higher among elderly people (e.g., Ceylan 2020; Remuzzi and Remuzzi 2020).

Understanding the substantial influence of demographic characteristics, particularly the size of the elderly population, population density, and access to handwashing facilities, is crucial to COVID-19 intervention and prevention strategies. Elderly population is a significant determinant in explaining the high mortality rates in Latin America and East and Southeast Asia but less important in Africa. Furthermore, population density might explain the dynamic of disease transmission, prevalence, and high mortality rates in many countries in Asia and Africa. In general, the Global South is more populated than the Global North. China and India account for almost one third of the world’s population, and many countries in the Global South are predominantly characterized by overcrowded residential environments. Therefore, although the governments and authorities imposed social distancing and stay-at-home restrictions to mitigate disease transmission, it might be difficult to successfully implement quarantine policies among overcrowded residents.

The normalized coefficients of importance revealed that HIV and TB prevalence are significant in explaining the high mortality prevalence especially across poorer Global South countries in sub-Saharan Africa and East Asia. In addition, the public health care infrastructure significantly affects preparedness and mitigation policies. The number of hospital beds is significant in treating critical cases and reducing a disease’s expected burden on a country’s health care system. Many Global South countries lack an adequate number of beds per 1,000 people and have no or limited resources to enhance the health care system. A considerable number of
countries in Africa have a low ratio of hospital beds to population, which is an indication of a low probability that these countries can respond to high demand for hospital beds arising from the COVID-19 pandemic. Moreover, the health care system in these countries faces entangled challenges and deficits such as poor health infrastructure, inequalities, low accessibility, and poor quality.

The geographic distribution of hospital beds per 1,000 population variable shows a negative association with mortality rate in some Latin American and Asian countries, and the absence of an adequate percentage of hospital beds in many African countries might explain the high current and estimated mortality rate, particularly in the central part of the continent. Many countries in the Global South spend about 6 percent or less of their GDP on the health care system, and the threshold target for health care expenditure has been estimated at 8.6 percent of GDP by 2030 for low-income and middle-income countries (Stenberg et al. 2017). To meet the third GDP by 2030 for low-income and middle-income countries, and 8.6 percent of GDP on health care. Most low-income countries, however, particularly in Africa, spend 6 percent or less of their GDP on the health care system. This rate of expenditure might not be enough to tackle the high incidence of COVID-19. Increasing health expenditure leads to greater numbers of professionals, physicians, nurses, and midwives. WHO recommends a minimum of 4.45 physicians per 1,000 people, but poor public health care system performance is associated with a ratio of less than 2.28 per 1,000 population. Many Global South countries lack the appropriate number of physicians, nurses, and midwives, so even if emergency resources are made available, having an inadequate number of professionals working in the health care system could represent a major challenge in the case of a pandemic. The high rate of infected critical cases is a challenge and poses threats to the robustness of the public health care system in many countries. Infected critical cases require ICU ventilator units, and the number of devices and volume of medical equipment in many Global South countries are inadequate. These cases could therefore be the trigger for effective planning of health infrastructure and services to avoid high mortality rates.

It is assumed that a high incidence of TB and HIV increases COVID-19 mortality rates in many Global South countries, mainly in Africa. TB impacts the lungs and hence infection with COVID-19 of a TB patient could result in a critical situation. HIV affects the immune system and results in its gradual and persistent failure and decline and can lead to acquired immunodeficiency syndrome. HIV results in intensified risk of life-threatening infection. The first defense against COVID-19 is the human immune system, and failure in this system results in increasing mortality rates. The rates of TB and HIV in developing countries, particularly in Africa, are high, and the health care systems in these countries could become overburdened by the demands of COVID-19 to the point where TB and HIV patients might not be able to access lifesaving treatment, increasing the mortality rate.

COVID-19 prevalence and mortality rates were estimated to be higher in middle-income countries (e.g., Iran, Brazil, India, Egypt, and Algeria) than in lower income countries (e.g., Mauritania, Mali, Sudan, Peru, and Mongolia). This finding could be explained by variations in health systems and medical equipment coverage, affordability, and accessibility. For instance, poor countries struggle to conduct sufficient COVID-19 tests. The number of confirmed cases varies significantly between countries due to the number of tests and the difference in epidemiological surveillance and detection capacities. Countries that conduct very few tests are most likely not testing widely enough to detect all cases. The transmission of the novel coronavirus can be mitigated and interrupted by early detection. The estimated mortality rates are high in central and northern Africa and eastern Asia, which indicates the need to conduct more tests to identify infected cases at early stages to allow authorities to respond appropriately to monitor and mitigate transmission of the disease.

It is noteworthy that the lower mortality rates in low-income countries, such as sub-Saharan countries, are likely to be associated with the testing approach and the underestimation of infected cases. For instance, and due to incomplete follow-up of COVID-19 deaths that did not receive a diagnosis, reported mortality percentages are likely to be underestimated. Accordingly, countries with a low number of tests per 1 million population also reported low rates of mortality. In contrast, the expanding level of testing in some countries might explain a large number of reported infected cases and fatalities. Accordingly, the spatial variations of mortality rates across the Global South are significantly associated with undetected and unreported cases. It is more likely that large numbers of infected cases have never been detected or reported because of inefficient medical infrastructure and health system deficits that lead to significant underestimation of the case-fatality ratio.

## Conclusion

The regions of the Global South are experiencing different levels of COVID-19 transmission, resulting in national geographic variations in total numbers of reported cases, incidence, and mortality rates during the first wave of the pandemic. Hence, modeling disease incidence and deaths across Global South countries is crucial to understanding and responding
The findings of this analysis reveal that strengthening health regimes, predominantly in hospitalization, critical intensive care, and number of physicians, could lead to substantial reductions in mortality rates in Global South countries. It is also important for public health authorities to manage medical care effectively and in a timely manner, as well as to direct the intensity and type of interventions needed to overcome the pandemic.

The associations between risks of COVID-19 mortality and demographic structure and health infrastructure present opportunities for targeted health promotion and preventive interventions in the Global South. Given the high burden of COVID-19 in these countries, it might be economically justified to implement intervention programs to protect high-risk population groups, particularly elderly populations. The incidence and transmuta-
tion of COVID-19 in the general population remains high, however, which poses an ethical dilemma in reliance only on high-risk strategies in these settings; countries in the sub-Saharan, Middle East, and North Africa regions might not have sufficient resources even to implement public health interventions in certain high-risk populations such as the elderly. Accordingly, addressing the wider sociodemographic determinants of the disease is therefore crucial to alleviate risk of mortality. For instance, in Yemen, Iran, and sub-Saharan African countries, while lockdown resulting from the pandemic has pushed millions to extreme poverty, the capacities of health systems are limited and thus they should be accompanied by massive amounts of aid and international support, particularly vaccine supplies, to avoid putting vulnerable populations at higher risk of mortality.

This research is limited by the absence of data on spatial variations of disease effects on a subnational scale and thus the effects of several local driving forces on mortalities that occur within smaller geographic zones might be muted at this higher level of analysis. Despite such limitations of data aggregation, however, this analysis has provided a valuable platform to plan further investigations around the epidemiology and management of COVID-19 in the Global South. As highlighted in the preceding limitations, there was a paucity of information on other important variables and ancillary predictors that can be used to model disease prevalence and mortality. Therefore, there is a need for more comprehensive country-specific data on COVID-19 and its determinants in Global South countries to account for the impact of these factors on disease prevalence predominately in low- and middle-income settings. The COVID-19 pandemic has led to unprecedented challenges, particularly across developing and poor countries. Therefore, geographic research on risk assessment as well as modeling the impacts of socioeconomics, demographic structure, and health care capacities on postdisease prevention syndrome in local societies is recommended. Likewise, more studies assessing COVID-19 morbidity and mortality are needed in less developed countries to ascertain the true prevalence of the disease and to accurately predict its future trends.

Acknowledgments

Authors’ contributions: Conceptualization, methodology, data curation, geospatial analysis, investigation, writing—original draft, S.M.; writing introduction section, data curation, writing—review and editing, A.A; data curation, writing—review and editing, M.A; data curation, writing—review and editing, E.R. All authors have read and agreed to the published version of the article.

Funding

Open Access funding was provided by the Qatar National Library.

ORCID

Shawky Mansour http://orcid.org/0000-0001-6969-9188
Ammar Abulibdeh http://orcid.org/0000-0002-0899-3655

Literature Cited


Idri, A., S. Mharki, and A. Abran. 2004. Validating and understanding software cost estimation models based on...


Science, Qatar University, Doha, Qatar. E-mail: aabulisdeh@qu.edu.qa. His research focuses on smart urban planning and design, sustainable built environment, sustainable transportation, and the water–energy–food nexus.

MOHAMMED ALAHMADI is an Associate Professor in the National Center for Remote Sensing Technology at King Abdulaziz City for Science and Technology (KACST), Riyadh 11442, Saudi Arabia. E-mail: mhalahmadi@kacst.edu.sa. He is an expert in modeling small area population data from satellite data. His research focuses on the application of machine learning, space-time modeling, and global environmental change.

ELNAZIR RAMADAN is an Assistant Professor in the Geography Department, Sultan Qaboos University, Muscat 123, Al Khoud, Oman. E-mail alnazir@squ.edu.om. His research interests cover issues related to land use planning and spatial development issues including sustainable urbanization with emphasis on the Global South. Further research interests include geospatial technologies application in spatial planning as well as urban governance and sustainability issues.