Heliyon 7 (2021) e06260

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification of risk factors contributing to COVID-19 incidence rates in Bangladesh: A GIS-based spatial modeling approach

Md. Hamidur Rahman [a,b,1], Niaz Mahmud Zafri [a,1,*], Fajle Rabbi Ashik [a], Md Waliullah [a], Asif Khan [a]

[a] Department of Urban and Regional Planning, Bangladesh University of Engineering and Technology (BUET), Dhaka 1000, Bangladesh
[b] Asian Disaster Preparedness Center (ADPC), Dhaka 1206, Bangladesh

ARTICLE INFO

Keywords:
Pandemic
Spatial variation
Spatial regression model (SRM)
Geographically weighted regression (GWR)
Demography
Built environment

ABSTRACT

Background: COVID-19 pandemic outbreak is an unprecedented shock throughout the world, which has generated a massive social, human, and economic crisis. Identification of risk factors is crucial to prevent the COVID-19 spread by taking appropriate countermeasures effectively. Therefore, this study aimed to identify the potential risk factors contributing to the COVID-19 incidence rates at the district-level in Bangladesh.
Method: Spatial regression methods were applied in this study to fulfill the aim. Data related to 28 demographic, economic, built environment, health, and facilities related factors were collected from secondary sources and analyzed to explain the spatial variability of this disease incidence. Three global (ordinary least squares (OLS), spatial lag model (SLM), and spatial error model (SEM)) and one local (geographically weighted regression (GWR)) regression models were developed in this study.
Results: The results of the models identified four factors: percentage of the urban population, monthly consumption, number of health workers, and distance from the capital city, as significant risk factors affecting the COVID-19 incidence rates in Bangladesh. Among the four developed models, the GWR model performed the best in explaining the variation of COVID-19 incidence rates across Bangladesh, with an R^2 value of 78.6%.
Conclusion: Findings and discussions from this research offer a better insight into the COVID-19 situation, which helped discuss policy implications to negotiate the future epidemic crisis. The primary policy response would be to decentralize the urban population and economic activities from and around the capital city, Dhaka, to create self-sufficient regions throughout the country, especially in the north-western region.

1. Introduction

Coronavirus disease (COVID-19) is an exceptionally infectious disease caused by the SARS-CoV-2 virus. The first human cases of COVID-19 were reported in Wuhan City, China, in December 2019 [1]. Within a few weeks, the outbreak of COVID-19 spread globally; thus, the World Health Organization (WHO) declared this a public health emergency of international concern (PHEIC) on 30 January 2020 and a pandemic on 11 March 2020. As of 05 February 2021, there have been about 105 million reported cases resulting in approximately 2.29 million deaths [2]. According to the World Bank, COVID-19 has triggered a global crisis like no other. The disease is leading to the deepest global recession since the Second World War. The baseline forecast envisions a 5.2% contraction in the global GDP in 2020— the deepest global recession in the last eight decades, despite unprecedented policy support [3]. About 1.6 billion informal workers, having little to no savings and no social protection access, lost 60% of their income. This pandemic will push 40–60 million people into extreme poverty [4].

Identification of possible demographic, economic, built environment, health, and service facilities related risk factors of infectious diseases is crucial at each phase of the epidemic to effectively prevent further spread through appropriate interventions [5]. Disadvantaged demographic pattern, economic, environmental, and health conditions have been established as potential determinants of infectious diseases (e.g., Tuberculosis, HIV, Pertussis, Pneumonia, SARS, Hand-foot-mouth, and Influenza) in general [5, 6, 7, 8, 9, 10]. These findings suggest that similar results might be obtainable for the newly emerged coronavirus disease. However, these

* Corresponding author.
E-mail address: zafri@urp.buet.ac.bd (N.M. Zafri).
[1] M. H. Rahman and N. M. Zafri contributed equally to this work as first authors.

https://doi.org/10.1016/j.heliyon.2021.e06260
Received 16 November 2020; Received in revised form 30 December 2020; Accepted 8 February 2021

indicators' impacts on the COVID-19 transmission and their magnitude are still under scrutiny and investigation. So far, several studies attempted to identify contributing risk factors, e.g., social [5, 11], economic [12], health [13], demographic [14, 15], environmental [16, 17, 18], which might affect the spread, morbidity rate, and mortality rate of the COVID-19 disease in the context of developed countries. Though very few studies are available in the context of developing countries [19, 20], results are still sparse. More studies need to be undertaken in different contexts, considering a wide variety of factors to get a complete scenario. Country contexts are critical as developing countries face major issues like large population, lack of infrastructure, social inequality, variety of employment, and inefficacy of public health measures, which might bring disparate outcomes in terms of virus incidence rates compared to developed countries [5].

Bangladesh, a highly dense developing country with many of the abovementioned attributes, is a perfect study unit for understanding the impact of contributing determinants of the COVID-19 outbreak. Besides, several important factors (e.g., built environment, service facilities related factors, and others) were not considered in the previous studies, which need to be explored to expand knowledge about the pandemic. Therefore, this study aimed to identify potential demographic, social, economic, built environment, health, and facilities-related determinants of the COVID-19 incidence rates at the district-level across Bangladesh. The findings of this study would help to prepare equitable public health prevention measures and guidelines for any future pandemic situation, especially in the context of developing countries.

2. COVID-19 situation in Bangladesh

The first three coronavirus cases were confirmed in Bangladesh on 8 March 2020 by the Institute of Epidemiology Disease Control and Research [21]. COVID-19 infected persons returned from Italy to join their family at their native places. Bangladesh reported its first coronavirus death on 18 March 2020 [22]. Following the initial detection of coronavirus cases in Dhaka, the virus transmitted from the capital city to other major administrative areas of Bangladesh rapidly. To intercept the disease's spread, the government declared a nationwide ‘lockdown’ from 26 March to 30 May 2020 [23], but unlike the developed countries, the enforcement of the lockdown was weak. To prevent viral transmission from higher infected to lower infected areas, the Bangladesh government took measures like the closure of educational institutions, declaration of a general holiday, restriction on religious gathering, suspension of commercial activities, closure of garments factories, prohibition of inter-country, inter-district, and intra-district travels [24]. To date, the education institutes are closed, but the government and non-government offices, as well as commercial and industrial entities, were opened after being shut down until 30 May 2020. Inter-city and intra-city travel restrictions were also lifted.

As of 31 January 2021, 0.535 million confirmed cases with 8,127 deaths (Infection Fatality Ratio is 1.52%), 0.48 million recoveries (89.6%), and 55 thousand (10.35%) active cases have been reported from 109 COVID-19 labs (63 laboratories within Dhaka, and 46 outside of Dhaka) as per Directorate General Health Services (DGHS) press release [21]. According to World Health Organization (WHO) [25], the highest death rate, 31.2%, was reported in 61–70 years old, and males represented 72% of reported confirmed cases in Bangladesh. In terms of geographic distribution, among the eight divisional units of Bangladesh, Dhaka division shows the highest 63.8% of total COVID-19 reported cases. As of 31 January 2021, Bangladesh overall attacked rate (AR) is 314 per 100,000, and all the 64 districts have reported COVID-19 confirmed cases. Figure 1 shows district wise COVID-19 incidence distribution, where Dhaka city has the highest attack rate, followed by Chattogram, Khulna, Sylhet, Rajshahi, and Barishal.

Like most other nations, the COVID-19 pandemic outbreak is an unprecedented shock to the economy of Bangladesh. The impact is particularly pronounced because of the country's heavy reliance on two sectors: ready-made garments (RMG) and remittances. According to Bangladesh Garment Manufacturers and Exporters Association (BGMEA), international buyers have either canceled or suspended $3.7 billion worth of shipments affecting three million workers as of 26 June 2020 [26]. Besides, the country's economy has inevitably been impacted by the reduction in the remittance flow, as 10 million migrant workers working in countries such as Saudi Arabia, the USA, Italy, the UK, and Malaysia, have been disadvantaged due to lockdowns and limited economic activities in these countries [27]. Financial organizations, small businesses, and startups are experiencing liquidity pressure as deposit growth and loan recovery declined due to COVID-19's impact on the country's economy as a whole. Based on the economic disruptions following the pandemic, the GDP growth forecast of Bangladesh by IMF, WB, and ADB has been revised down from 7.8%-8.2% to a range between 2.0%-3.8% for FY′20. They project export to slow down by 15.4% and import to fall by 11.8% [28].

3. Data and methods

3.1. Data collection and preparation

Institute of Epidemiology, Disease Control and Research (IEDCR) has monitored the spread of the pandemic and updated the database of COVID-19 daily at district and city level across Bangladesh. For this study, data about the number of positive COVID-19 cases at the district-level across Bangladesh from 8 March 2020 (first known instances in Bangladesh) to 1 August 2020 were considered, retrieved from the IECDR data portal [21]. COVID-19 cases per 10,000 population per district were considered as the dependent variable for modeling and interpretation purposes and termed “COVID-19 incidence rates” in this study.

A total of 28 demographic, economic, built environment, health facilities, and community facilities related factors were considered explanatory variables for the model development. Most of the data of relevant factors were compiled from different databases of Bangladesh government, including the Bangladesh Bureau of Statistics (BBS), the centralized official bureau in Bangladesh for collecting statistics on demography, economy, and other facts of the country. Few factors were derived through spatial analysis using ArcGIS Pro 2.4 software. Most of the variables have been normalized either by taking a thousand population per unit or taking a percentage of the total distribution. A detailed description of the variables and their sources is shown in Table 1. After data collection and preparation, district-level information was tabulated as an attribute table in ArcGIS.

3.2. The use of GIS and spatial modeling approach

The use of geospatial and statistical tools is critical to explore the association between COVID-19 incidence and its contributing factors as it is a process that occurs in geographical space [29]. Traditional statistical approaches used in epidemiological studies, e.g., factor analysis [30], principal component analysis [31], cluster analysis [32], regression analysis [33] fail to take into account the spatial dependency and autocorrelation in parameters estimation. To address this functional lag, the spatial regression model (SRM), i.e., spatial lag model (SLM), spatial error model (SEM), and geographically weighted regression (GWR) have been widely used in epidemiological studies [14]. Besides, geographic information system (GIS) is an essential tool that can assist in the process of combating a pandemic as well as improve the quality of care through examining the spatial distribution of infectious diseases [29, 34, 35].

In this study, data of explanatory variables are inherently spatial. Hence, it is plausible that data would be positively spatially dependent, which means areas located nearby tend to be more similar than those separated by great distances [36]. As OLS regression cannot take spatial dependence into account due to the assumption of homogeneity and spatial non-variability [37, 38], we used SEM and SLM along with OLS.

Nevertheless, global models still have limitations as they cannot account for a spatial non-stationarity issue. The spatial non-stationarity issue explains how the dependent and independent variables might vary over space [14, 39]. Thus, as a local regression model, GWR was used to address the spatial non-stationarity issue. A detailed description of this model has been avoided to focus more on discussions than methodological exercises. Information on these models is available in the studies of Mollalo, Vahedi [17] and Sannigrahi, Pilla [14].

3.3. Model development

At first, for developing multivariate models, a univariate analysis was performed to identify the potential explanatory variables from the collected data of many factors. Here, OLS models were developed in ArcGIS to explain each variable's impact on the COVID-19 incidence rates individually. Factors found insignificant in this stage were excluded for further consideration as explanatory variables. From these selected explanatory variables, the stepwise forward procedure was applied to eliminate non-significant explanatory variables and develop an overall

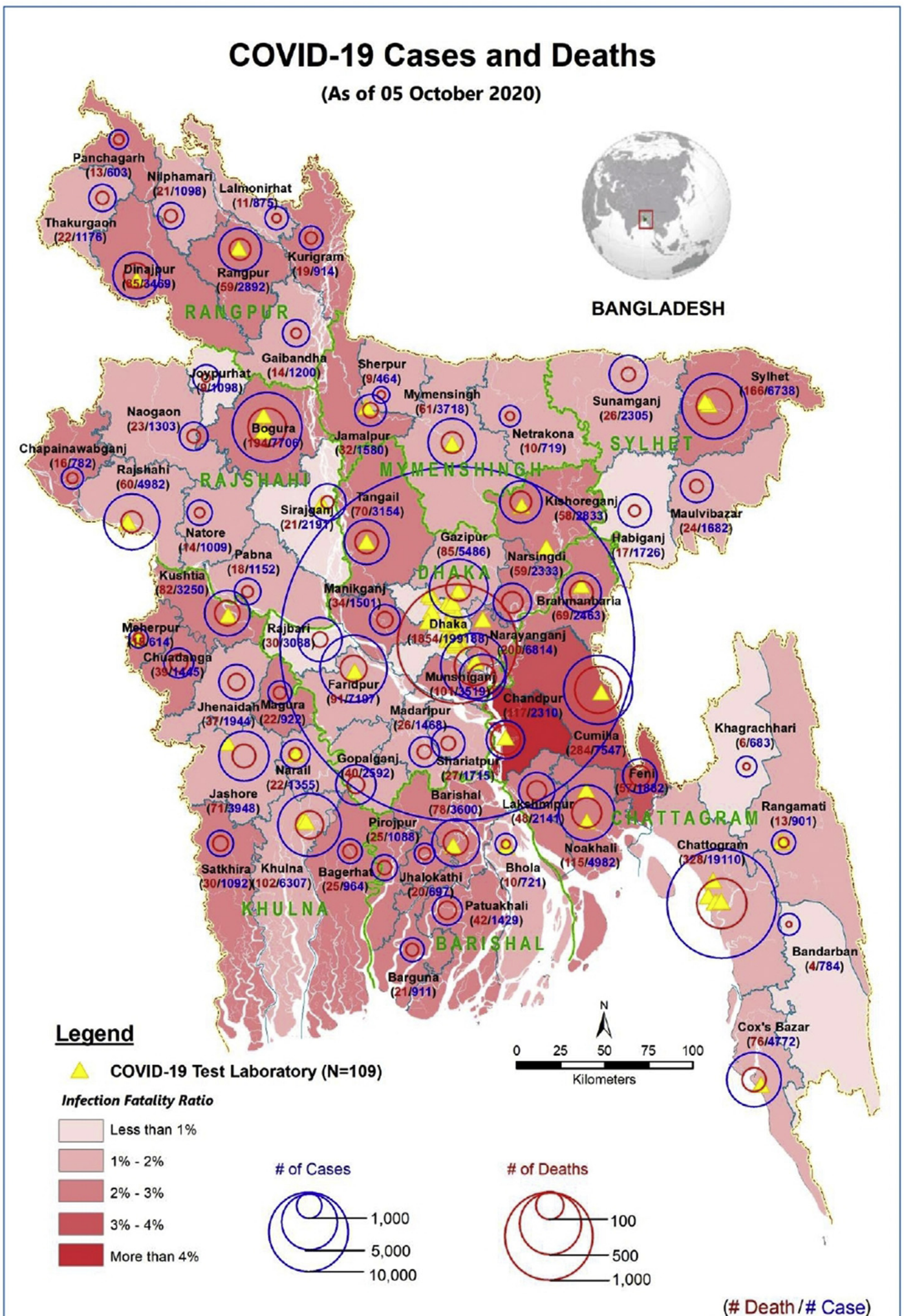

Figure 1. Map showing the distribution of COVID-19 cases and fatality, 08 March-05 October 2020, Bangladesh [25].

Table 1. Collected factors used in this study together with definitions and sources.

Factor	Description	Source
Demographic factors		
(1) Total population	The total population lived in a district ('0000)	(1,2) Population Monograph 2015: Vol.07
(2) Population density	Population living per sq. km. of the district area	
(3) Urban population percentage	% of urban population in a district	(3,4) Population Monograph 2015:Vol.06, BBS
(4) Dependency ratio	The ratio of population aged 65 + per 1,000 population aged 15-64	
(5) Slum population percentage	% of the population living in slum areas in a district	(5) Census of Slum Areas and Floating Population 2014: Vol-06, BBS
(6) Rented tenancy percentage	% of households occupied by rents in a district	(6) Population Monograph 2015:Vol.10, BBS
(7) Internal migrant population	The total population who moved across or within the districts ('0000)	(7) Population and Housing Census 2011, National Report: Volume-4, BBS
Economic factors		
(8) Poverty rate	% of people living below the upper poverty line as a share of the total district population	(8,9,10)) Household Income and Expenditure Survey (HIES)-2016, BBS
(9) Income inequality	Income inequality measured by the Gini coefficient for each district	
(10) Number of the economic units	Total economic establishments within a district boundary	
(11) Refined activity rate	The ratio of economically active population to the population aged over 14 years in a district	(11) Population Monograph 2015-Vol.11, BBS
(12) Monthly consumption	Average monthly consumption (on food and non-food consumption as well as inventory of durable goods) per households in BDT in a district ('000)	(12) Household Income and Expenditure Survey (HIES)-2016, BBS
Healthcare facilities related factors		
(13) Number of community clinics	Total number of community clinic in a district	
(14) Number of upazila health complexes	Total number of upazila health complex in a district	
(15) Number of health workers	Total number of health workers, i.e., doctors, nurses in a district ('000)	
Built environment related factors		
(16) Urban land area	Total urban land (in sq. km.) in a district	(16,17,18) Zilla Statistics, BBS
(17) Rural land area	Total rural land (in sq. km.) in a district	
(18) Road density	Road length (km) per sq. km. of the district area	
(19) Distance from the capital	Centroidal distance between a district from the capital city, Dhaka (in kilometer)	(19,20) Author's calculation through GIS
(20) Distance from the divisional headquarter	Centroidal distance between a district and the corresponding divisional headquarter of that district (in kilometer)	
Community facilities related factors		
(21) Number of primary schools	Total number of primary schools in a district	(21,22,23,24,25) Local Govt. Engineering Division (LGED), 2018
(22) Number of secondary schools	Total number of secondary schools in a district	
(23) Number of colleges	Total number of colleges in a district	
(24) Number of growth centers	Total number of growth centers in a district	
(25) Number of rural markets	Total number of rural markets in a district	
(26) Number of religious establishments	Total number of religious establishments, i.e., mosque, temple in a district	(26,27,28) Zilla Statistics, BBS
(27) Number of transit stations	Total number of transit stations, i.e., bus stand, lunch terminal, rail station in a district	
(28) Number of police stations	Total number of police stations in a district	

global multivariate OLS model with the best fit statistics. Then, Pearson's correlation analysis was conducted to examine the correlations between the significant factors in univariate analysis. For detecting multi-collinearity in the model, the Variance Inflation Factor (VIF) was used, and therefore, the uncorrelated factors were selected as the input of the final OLS regression model. After that, SLM, SEM, and GWR were developed using the final OLS model's significant explanatory variables for comparison purposes. Two global models (SLM and SEM) were developed in GeoDa 1.14 software. Based on the first-order Queen's contiguity, the weight matrix was generated for developing the SLM and SEM.

On the other hand, the local model (GWR) was run in ArcGIS software. *Adaptive* kernel type and *AICc* bandwidth were selected to run the GWR model. Finally, The R^2 and *AICc* values of the four developed models were used to compare the models' performances in explaining COVID-19 incidence rates across Bangladesh.

4. Results

The results of the univariate analysis are presented in Table 2. Out of 28 factors, 17 were found statistically significant and considered explanatory variables for developing a multivariate model later. All the demographic factors were found significant; whereas, most of the community facilities related factors were insignificant. From Table 2, it is also clear that most of the demographic and health-related factors have comparatively higher R^2 values, which means these factors could explain a good portion of the variation in the COVID-19 incidence rates across Bangladesh. On the other hand, a relatively lower R^2 was found for other

Table 2. Results of univariate analysis.

Factor	Intercept	Coefficient	Std. Error	*p*-value	R^2
Demographic factors					
Total population	2.492	0.023	0.004	0.000*	0.366
Population density	2.658	0.004	0.001	0.000*	0.507
Urban population percentage	-0.973	0.476	0.054	0.000*	0.561
Dependency ratio	18.684	-0.163	0.074	0.032**	0.072
Slum population percentage	5.035	2.791	0.473	0.000*	0.359
Rented tenancy percentage	4.348	0.462	0.046	0.000*	0.519
Internal migrant population	6.166	0.069	0.008	0.000*	0.529
Economic factors					
Poverty rate	12.294	-0.168	0.050	0.001*	0.153
Income inequality	11.316	-13.007	21.749	0.552	0.006
Number of economic units	2.675	0.040	0.007	0.000*	0.344
Refined activity rate	-15.887	0.803	0.191	0.000*	0.221
Monthly consumption	-5.766	0.928	0.189	0.000*	0.279
Healthcare facilities related factors					
Number of community clinics	9.453	-0.013	0.015	0.389	0.012
Number of upazilas health complexes	8.295	-0.029	0.038	0.454	0.009
Number of health workers	5.068	1.253	0.149	0.000*	0.533
Built environment related factors					
Urban land area	4.928	0.020	0.010	0.049**	0.060
Rural land area	7.674	-0.001	0.001	0.986	0.000
Road density	6.464	0.482	0.615	0.435	0.010
Distance from the capital	12.315	-0.031	0.010	0.004*	0.127
Distance from the divisional headquarter	11.225	-0.049	0.017	0.006*	0.114
Community facilities related factors					
Number of primary schools	5.825	0.002	0.002	0.356	0.014
Number of secondary schools	5.929	0.006	0.006	0.313	0.016
Number of colleges	6.030	0.036	0.034	0.284	0.018
Number of growth centers	5.411	0.070	0.064	0.274	0.019
Number of rural markets	8.450	-0.003	0.007	0.642	0.004
Number of religious establishments	5.189	0.001	0.000	0.134	0.036
Number of transit stations	4.879	0.053	0.018	0.005*	0.123
Number of police stations	4.122	0.162	0.044	0.001*	0.179

* Significant at 99% confidence level.
** Significant at 95% confidence level.

factors. Therefore, variation in COVID-19 incidence rates could mostly be described by demographic and health-related factors.

After feature selection and correlation analysis, among the 17 explanatory variables, only four variables were included in the final global multivariate OLS model (Figure 2). Several factors, which had high R^2 values and were strongly significant in univariate analysis, were not considered for the model development to diminish multicollinearity in the model. Those factors were: total population, population density, rented tenancy percentage, internal migrant population, and the number of an economic unit. The final model included only four significant factors— urban population percentage, monthly consumption, number of health workers, and distance from the capital (Table 3).

The model has relatively low multicollinearity since the highest VIF value among the four factors (VIF = 2.9) is far less than the threshold of 7.5. The association's direction suggested that three factors positively associate with the COVID-19 incidence rates except for distance from the capital factor. This model was found to be statistically significant and has an R^2 value of 0.673. This value means that about 67.3% of the COVID-19 incidence rates across Bangladesh are associated with the model's four factors. The rest of the 32.7% incidence rates are caused by unknown factors to the model and probably for the local variations, which could not be captured by the global OLS model.

Later, SLM and SEM models were developed by incorporating spatial dependence among the variables to improve the results of the overall OLS model (Table 4). Autoregressive lag coefficients (Rho and Lamda) of the models were strongly statistically significant at a 99% confidence level. Both SLM and SEM have higher R^2 values and lower *AICc* values than the OLS model (Table 5), which indicates SLM and SEM performed better than the OLS model. However, the performance of modeling the COVID-19 incidence rates in Bangladesh might be further improved if the model could be developed on a local scale instead of a global scale.

GWR was used to model the COVID-19 incidence rates on a local scale. From Table 5, the R^2 value was found to be the highest for the SEM model among the global models. This value increased from 0.722 in the SEM to 0.786 in the GWR model. Therefore, it is evident that the GWR model could explain 78.6% of the variations of COVID-19 incidence rates across Bangladesh. In addition, the *AICc* value was also found the lowest in the GWR model (*AICc* = 340.49) compared to the global models, indicating that the GWR model is the most effective in this context.

The spatial distributions of the coefficient values of the GWR model are presented in Figure 3. As seen in Figure 3, monthly consumption and the number of health workers factors demonstrated nearly similar patterns; the opposite pattern was found for urban population percentage. For the eastern and southern regions of Bangladesh, monthly consumption and the number of health workers were found as influential factors, and urban population percentage was found as a weak factor in explaining the COVID-19 incidence rates. On the other hand, for the north-western regions of the country, the opposite findings were

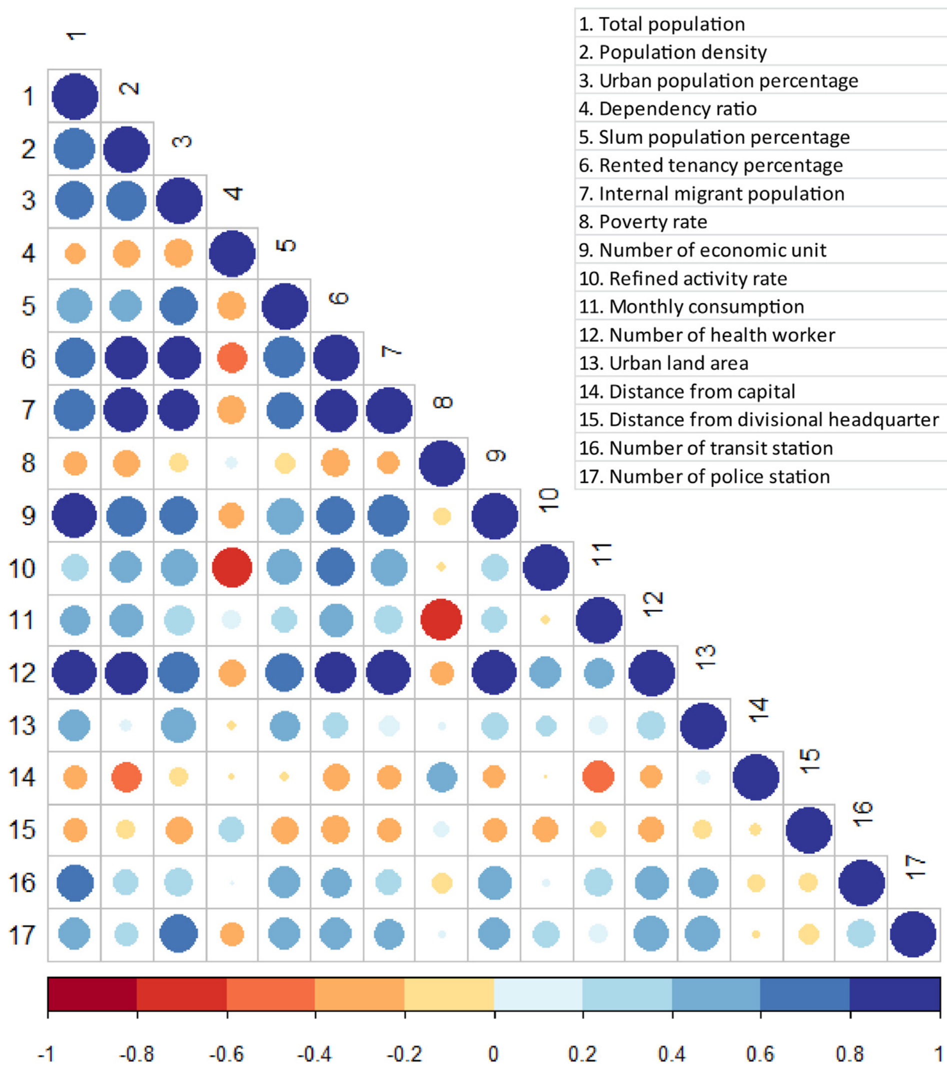

Figure 2. Results of the correlation analysis.

Table 3. Results of the final multivariate OLS model.

Factors	Coefficient	Std. error	*p*-value	VIF
Intercept	-0.966	2.890	0.738	
Urban population percentage	0.278	0.080	0.000*	2.857
Monthly consumption	0.314	0.157	0.046**	1.450
Number of health workers	0.468	0.217	0.035**	2.903
Distance from the capital	-0.013	0.007	0.079***	1.263
Model statistics: F (4, 59) = 30.401, p = 0.000, R^2 = 0. 673				

* Significant at 99% confidence level.
** Significant at 95% confidence level.
*** Significant at 90% confidence level.

observed. Furthermore, distance from the capital was an influential factor in the country's central parts. It was less influential in the south-eastern region of the country. The discussion section explains why and how these factors better explain COVID-19 incidence in a developing country.

The results of mapping local R^2 values of the GWR model are demonstrated in Figure 4. Though decent local R^2 values were found for all districts, the districts located in the southern regions of the country have comparatively lower local R^2 values than northern regions of the country, indicating a decent prediction of the model across the country, especially in northern districts. The reason behind the slight variation in the local R^2 value might be unique topographical differences between the two regions. In the southern parts of Bangladesh, coastal districts are located, which are mostly isolated by rivers from their surrounding districts. In addition, southern-east parts of Bangladesh are hilly areas where communication from one district to another district is very difficult. Whereas, central, northern, and north-western parts of the country are relatively flat land areas. Therefore, communication among them is relatively easier. As the GWR model estimated the local R^2 value based on the surrounding districts, therefore, for the southern part of the country,

Table 4. Results of the SLM and SEM models.

Variable	Coefficient		Std. error		p-value	
	SLM	SEM	SLM	SEM	SLM	SEM
Intercept	-2.309	-1.558	2.583	3.251	0.371	0.631
Urban population percentage	0.237	0.299	0.074	0.072	0.000*	0.000*
Monthly consumption	0.181	0.318	0.149	0.164	0.033**	0.037**
Number of health workers	0.565	0.346	0.202	0.182	0.005*	0.057***
Distance from the capital	-0.006	-0.011	0.006	0.009	0.025**	0.015**
Rho	0.356		0.113		0.001*	
Lamda		0.466		0.135		0.000*

* Significant at 99% confidence level.
** Significant at 95% confidence level.
*** Significant at 90% confidence level.

surrounding districts cannot influence much on the COVID-19 incidence rates of the concerned district due to poor communication. Therefore, it is likely that for these reasons the GWR model performed slightly inferior in that region compared to other regions.

5. Discussion

The percentage of the urban population of districts was positively related to COVID-19 incidence rates, which is similar to the findings of Hamidi, Sabouri [15], where researchers found that a high urban population leads to increase movement and activities of people in a high-density urban area. The higher the population density, the more likely it is for an infectee to contact an infector [40]. This study also found population density as a significant variable in univariate analysis. However, it could not be considered for the final model due to the multicollinearity issue. In addition to that, the impact of the urban population percentage on the COVID-19 incidence rates was found lower (lower coefficient value in GWR model) in the north-western region (e.g., Rajshahi, Rangpur, and Mymensingh division) of Bangladesh compared to the south, central, and south-eastern regions (Figure 3). Here, the geographical context of the south-central-eastern regions of Bangladesh suggests that these regions cover three major urban areas. The first one is Dhaka megacity, which comprises 38% (12 million people) of the country's total urban population. Next is Chattogram, the second-largest city located in the south-eastern region of Bangladesh, which has one-third population of Dhaka. Lastly, Khulna, the third-largest city of the country, is also located in the southern region [41]. All of these three cities are densely populated and have greater economic and administrative agglomeration than other parts of the country. On the other hand, the north-western region's major urban areas are not as vibrant as the mentioned three cities as they are far behind in employment opportunities, commercial activities, and industrial activities [41]. As a result, COVID-19 incidence rates were less influenced by the urban population in the north-western region than the country's other regions.

In this situation, planned and decentralized urbanization needs to be ensured and physical and social infrastructure need to be better distributed across the country so that people do not concentrate in a few districts, raising their densities beyond sustainable limits. For decentralized urban development, existing cities of the north-western region of Bangladesh need to be prioritized as alterative urban areas of Dhaka and Chattogram to reduce pressure from these two cities.

The monthly consumption factor positively influenced COVID-19 incidence rates. Households having higher monthly consumption tend to purchase goods from commercially vibrant places, which increases the potential of being affected by COVID-19 [40]. Furthermore, higher consumption is associated with higher income levels, lower poverty, and higher employment rates. Therefore, income and employment-related activities trigger frequent travel and physical contact with people, increasing the risk of transmission of COVID-19 [42]. Another explanation behind this could be that low-income people might have less access to testing facilities, resulting in the underreporting of the positive COVID-19 cases in low-income areas [11, 43]. In addition to that, monthly consumption had a higher impact on COVID-19 incidence rates in Bangladesh's north-western region than in other regions (Figure 3). The north-western region of the country is lagging behind than other regions due to higher poverty and inequality [44]. This region suffers from severe "Monga", referred to as a seasonal phenomenon of poverty and hunger, during the time period of September–November and March–April [45]. So, after the March–April "Monga", people are compelled to take part in economic activities to secure their needs ahead of the upcoming "Monga" which might lead to the higher impact of monthly consumption on COVID-19 incidence rate in this particular region.

As a recommendation, economic activities need to be diversified and dispersed around the country so that high-income people and employment opportunities are not concentrated in a few districts. A large portion of the north-western region's poor working group migrates to Dhaka for work due to a lack of employment in their native place [46]. They are mostly engaged in informal jobs in Dhaka. In the period leading from the first detection of the COVID-19 case in Bangladesh, quite a few people lost their jobs while many others experienced pay cuts [47]. As a result, many households left the capital, Dhaka, facing affordability issues, and they mostly relocated to their village homes [48, 49]. This is the time to seize the opportunity to create employment for these people by providing them monetary support to engage in agricultural activities and establish industries, which can be based in remote areas, especially the north-western region. This will ensure that these people will be less likely to re-migrate to the capital; instead, others living in Dhaka and other major cities might be attracted to return to their village homes.

The number of health workers positively influenced the COVID-19 incidence rates as hospitals became the hotspots for COVID-19 circulation. This result was found to be consistent with the findings of Mollalo,

Table 5. Measures of goodness-of-fit for OLS, SEM, SLM, and GWR models.

Criterion	OLS	SLM	SEM	GWR
R^2	0.673	0.717	0.722	0.786
AICc	363.94	355.18	353.63	340.49

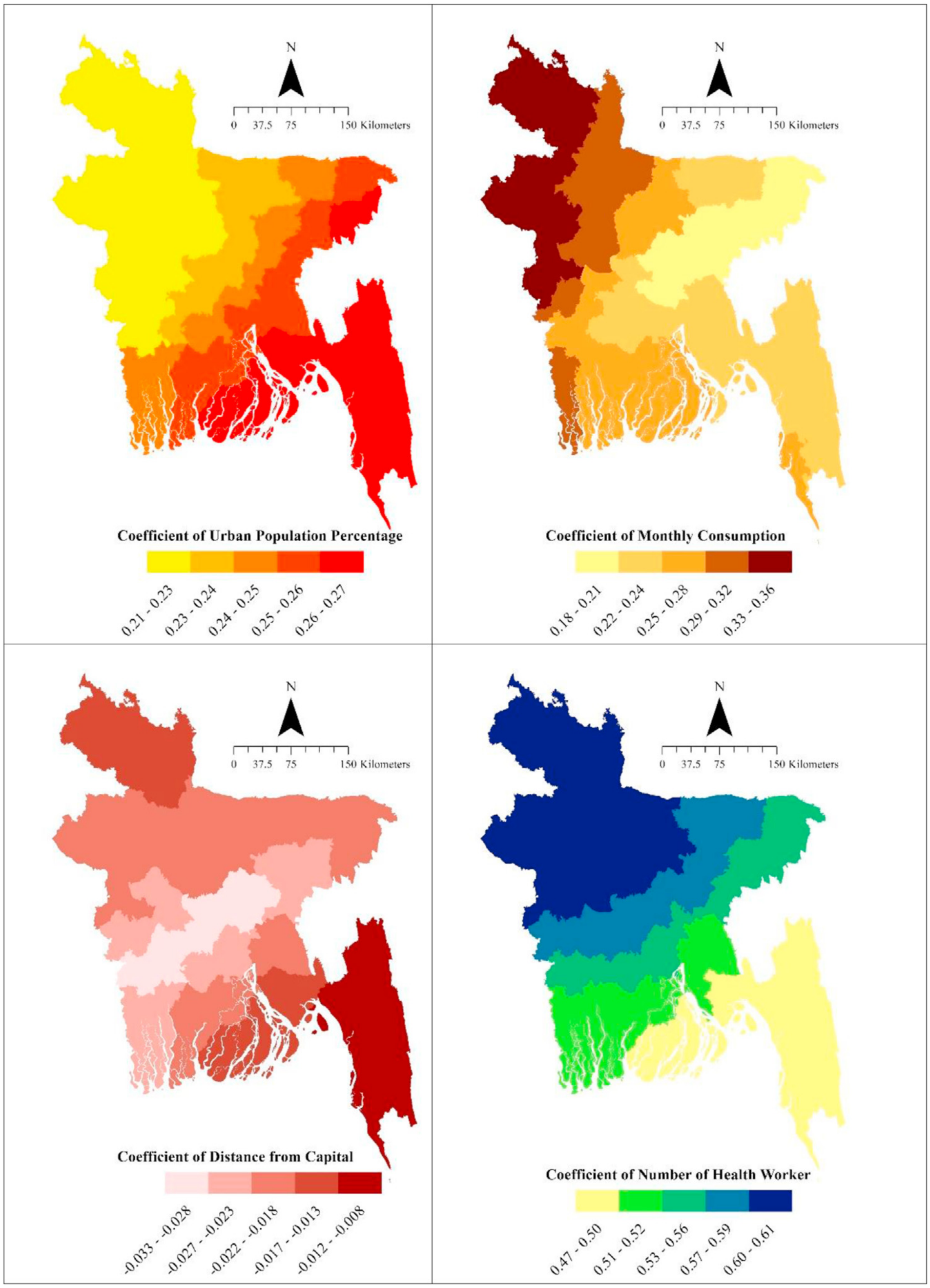

Figure 3. Spatial distribution of the coefficient values of urban population percentage, monthly consumption, number of health workers, and distance from the capital in describing COVID-19 incidence rates using GWR model.

Vahedi [17]. From correlation analysis, it can also be observed that the number of health workers factor had a strong positive relationship with total population and population density (Figure 2). This direction of correlation suggested that areas with a larger population with higher population density have comparatively more health facilities and, consequently, have a higher number of health workers and a number of COVID-19 testing centers. Therefore, the number of positive COVID-19 cases was also more in those areas. In addition, this factor was found to have a more substantial influence on COVID-19 incidence rates in the north-western region of Bangladesh than in other regions (Figure 3). There was a severe shortage of personal protection equipment (PPE) in Bangladesh at the beginning of the pandemic. Therefore, a significant percentage (25%) of front-line health workers were obliged to tackle COVID-19 without any protection, and thereby, a large number of health workers became infected by COVID-19 [50]. The infected health professionals are likely to have passed on the disease to others, including patients visiting the hospitals for non-COVID-19 related issues and people accompanying them. Furthermore, PPEs were distributed in the hospitals of important urban areas first. Therefore, it might be assumed that hospitals in the north-western region received PPE later. Due to this, COVID-19 might have spread from hospitals through health workers in this region.

In a situation like this, the supply-chain system must be uninterrupted so that proper protective equipment can be delivered to the health workers on a priority basis. The demand for health services escalates during pandemic times. Steps need to be taken to reduce the pressure on existing hospitals and prevent them from becoming COVID-19 hotspots. A few places in the major cities can be earmarked for erecting temporary hospitals to tackle a pandemic like the COVID-19. These places should

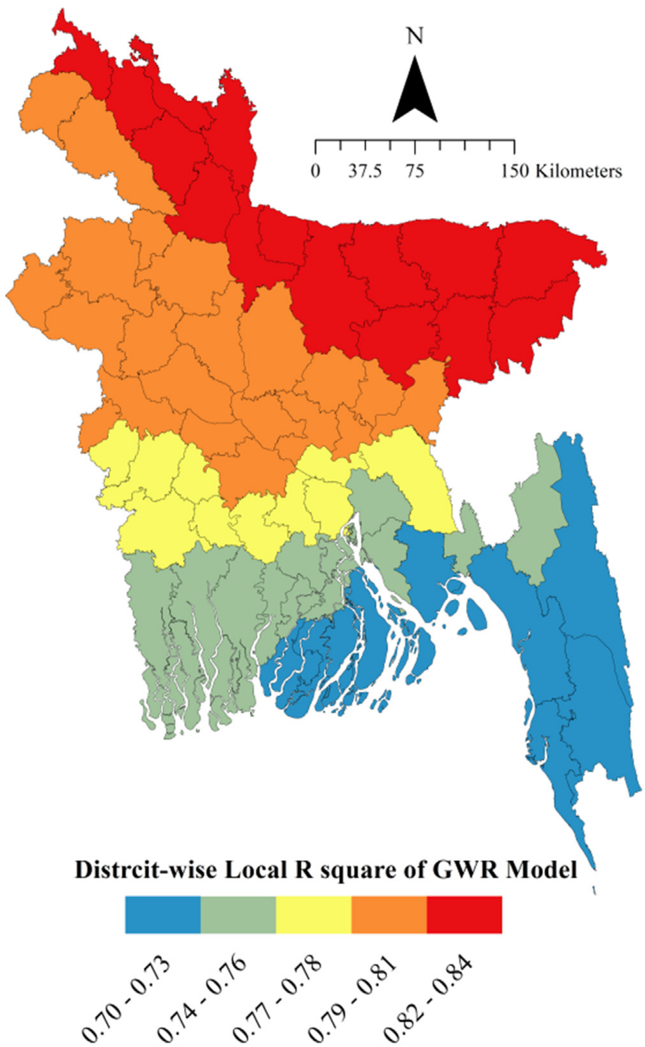

Figure 4. Spatial distribution of the local R^2 values of GWR model.

have good transport access, and they can be used as recreational open spaces at other times.

Distance from the capital was found inversely related to COVID-19 incidence rates. Dhaka, the capital and a megacity of Bangladesh, is the epicenter of the country's pandemic and has the most infected population (Figure 1). In addition, Dhaka is the central commercial and administrative hub of the country; hence, it attracts people from all over Bangladesh, resulting in many trips to and from Dhaka. Since Dhaka is the economic hub of Bangladesh, a lot of people live in surrounding districts to save on housing costs and commute to the city daily. During this pandemic, public transport was under restriction across the country. People residing in districts in closer proximity to Dhaka could somehow manage to travel to this district more frequently than districts located farther. These might be the reasons behind the result— the shorter the distance of a district from the capital, the higher the likelihood of transmission of COVID-19. Moreover, correlation analysis suggested that distance from the capital was negatively correlated with population density while positively correlated with poverty rate. The more the districts are far away from the capital city, the greater is their potential to have low density and higher poverty incidence which negatively affected COVID-19 incidence rates in Bangladesh (Figure 2). Furthermore, this factor was less influential in the south-eastern region, southern region, and the north-eastern region of Bangladesh compared to other regions. The reason behind that is the south-eastern region is a hilly area, and the south region comprises a large number of islands where water transportation is the primary mode for communication. Therefore, accessibility between the capital city and these locations is not as good as with other regions.

Additionally, travel between a lagged north-western region and Dhaka might be less required during the pandemic. For these reasons, fewer viral transmissions occurred in those regions from Dhaka, and fewer COVID-19 incidences were reported there. From literature, connectivity was also found more important than density in 913 US metropolitan counties [15]. The policy response of diversifying and dispersing economic activities across Bangladesh should be able to take care of this issue and reduce pressure on the capital city. This would help to create self-sufficient regions and might reduce the demand for inter-region travel during critical situations.

This study showed that the impacts of most of the community facilities, e.g., primary school, secondary school, college, growth center, rural market, and religious establishment on COVID-19 incidence rates were insignificant as these facilities were shut down in the initial stage of COVID-19 pandemic. However, the univariate analysis results show that the number of transit stations and the number of police station factors profoundly influenced increasing COVID-19 incidence rates. Though transit stations were also controlled through lockdown measures, there was a lax during the festive periods, and garment factories were hastily opened and immediately closed, which resulted in the movement of a large number of people [51]. In addition, emergency activities, such as health care, food delivery, and necessary regular goods marketing were out of the scope of lockdown measures and involved the use of transit stations. As a result, transit stations might have become crowded, which paved the way to transmit COVID-19. Police had the responsibility to implement lockdown and social distancing at the field level during the pandemic. Enough protection equipment was not available for the police personnel to carry out their duties safely. Moreover, police officers' overcrowding travel in a single police van made them more vulnerable to the virus [52]. The design of transit stations needs to take the physical distance factor into account. The design should discourage the clustering of people and ensure good natural air circulation. Like the health professionals, police personnel and all other front-line workers should be given priority for the distribution of PPE after ensuring an adequate supply of this equipment through an efficient supply-chain system.

6. Limitation of the study

One of the limitations of our study was the lack of appropriate data for some variables. Due to the unavailability of individual or community level data, it is not logical to draw inferences at the individual or community level. Another reservation is the COVID-19 testing center's spatial availability. There are few opportunities to test COVID-19 for people living in remote areas in Bangladesh. Besides, there is a tendency among people not to test COVID-19 even though they have symptoms. Therefore, there might be an underestimation of COVID-19 cases. Furthermore, the influence of lockdown and other containment measures on COVID-19 incidence rates were not considered in this study. There is likely to be some variations in lockdown related policies and their implementation efficiency within a district and between districts. It might play an essential role at the district-level to control the COVID-19 incidence rates, but analyzing this influence was out of this research scope.

7. Conclusion

Identification of possible virus transmission and spread determinants is crucial, especially for coronavirus disease (COVID-19), which brought unprecedented shock globally. This study aimed to identify potential risk factors contributing to the district-level COVID-19 incidence rates across Bangladesh. To fulfill the aim, three global (OLS, SLM, and SEM) and one local (GWR) models were developed in this study to identify potential demographic, economic, built environment, health, and facilities related factors affecting the COVID-19 incidence rates.

The models' results showed four factors— urban population percentage, monthly consumption, number of health workers, and distance from the capital— were statistically significant. The R^2 value was found to be 0.673 for the overall OLS model, which increased to 0.78 in the GWR model. The findings of the study showed that the COVID-19 incidence rates increased with an increase in urban population percentage, monthly consumption, and the number of health workers within a district. At the same time, the incidence rates decreased with an increase in the distance between the capital city with districts. Discussions on the findings reveal that higher level dependency on and concentration of economic as well as industrial activities within a few districts, ineffective supply-chain system, lack of self-sufficient regions, poverty and inequality, especially in the north-western region of the country, were some of the main causes which led to higher COVID-19 incidence rates.

The recommended policy response suggests that increasing dependency on the capital city should be lessened by diversifying and decentralizing economic activities across Bangladesh, especially the north-western region of Bangladesh, which needs to be prioritized as alternative urban centers for Dhaka and Chattogram. This would help create self-sufficient regions and might reduce the demand for inter-region travel during critical situations and, consequently, help control any pandemic like the COVID-19 in the future. Theoretical investigations and empirical observations from this research offer an alternate view of the joint importance of health and non-health determinants, which will help planners and local governments for effective policy development to tackle future epidemic crises. Future studies incorporating individual or community level data having temporal consistency with COVID-19 transmission time could produce better directives analyzing risk factors at a lower geographic unit and finding solutions for different socio-economic groups.

Declarations

Author contribution statement

M. H. Rahman, N. M. Zafri: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

F. R. Ashik, M. Waliullah: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

A. Khan: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The research work is facilitated by the support of Bangladesh University of Engineering and Technology (BUET) through their research lab facilities and infrastructures. Hence, the authors are happy to acknowledge that support. In addition, the authors also like to thank Madiha Chowdhury (URP'15, BUET) for proofreading the manuscript.

References

[1] WHO, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), World Health Organization, 2020.
[2] WHO. WHO Coronavirus Disease (COVID-19) Dashboard. World Health Organization2020.
[3] WorldBank, Global Economic Prospects, World Bank Publications, Washington, DC, 2020.
[4] UN, BRIEF#2: PUTTING THE UN FRAMEWORK FOR SOCIO-ECONOMIC RESPONSE TO COVID-19 INTO ACTION: INSIGHTS, United Nation, 2020.
[5] S. Khalatbari-Soltani, R.C. Cumming, C. Delpierre, M. Kelly-Irving, Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards, J. Epidemiol. Community Health 74 (8) (2020) 620–623.
[6] R. Bluhm, M. Pinkovskiy, The spread of COVID-19 and the BCG vaccine: a natural experiment in reunified Germany, FRB New York Staff Rep. (926) (2020).
[7] T. Bärnighausen, V. Hosegood, I.M. Timaeus, M.L. Newell, The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa, AIDS (London, England) 21 (Suppl 7) (2007) S29–38.
[8] X. Huang, S. Lambert, C. Lau, R.J. Soares Magalhaes, J. Marquess, M. Rajmokan, et al., Assessing the social and environmental determinants of pertussis epidemics in Queensland, Australia: a Bayesian spatio-temporal analysis, Epidemiol. Infect. 145 (6) (2017) 1221–1230.
[9] B.M. Farr, C.L. Bartlett, J. Wadsworth, D.L. Miller, Risk factors for community-acquired pneumonia diagnosed upon hospital admission. British Thoracic Society Pneumonia Study Group, Respir. Med. 94 (10) (2000) 954–963.
[10] L.-Q. Fang, S.J. de Vlas, D. Feng, S. Liang, Y.-F. Xu, J.-P. Zhou, et al., Geographical spread of SARS in mainland China, Trop. Med. Int. Health : TM & IH 14 (Suppl 1) (2009) 14–20.
[11] J. Corburn, D. Vlahov, B. Mberu, L. Riley, W.T. Caiaffa, S.F. Rashid, et al., Slum health: arresting COVID-19 and improving well-being in urban informal settlements, J. Urban Health (2020) 1–10.
[12] A. Atkeson, What Will Be the Economic Impact of Covid-19 in the Us? Rough Estimates of Disease scenarios, National Bureau of Economic Research, 2020, pp. 898–2937. Report No.
[13] Q. Zhao, M. Meng, R. Kumar, Y. Wu, J. Huang, N. Lian, et al., The impact of COPD and smoking history on the severity of COVID-19: a systemic review and meta-analysis, J. Med. Virol. (2020).
[14] S. Sannigrahi, F. Pilla, B. Basu, A.S. Basu, A. Molter, Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach, Sustainable Cities Soc. 62 (2020) 102418.
[15] S. Hamidi, S. Sabouri, R. Ewing, Does density aggravate the COVID-19 pandemic? J. Am. Plann. Assoc. (2020) 1–15.
[16] J. Liu, J. Zhou, J. Yao, X. Zhang, L. Li, X. Xu, et al., Impact of meteorological factors on the COVID-19 transmission: a multi-city study in China, Sci. Total Environ. (2020) 138513.
[17] A. Mollalo, B. Vahedi, K.M. Rivera, GIS-based spatial modeling of COVID-19 incidence rate in the continental United States, Sci. Total Environ. (2020) 138884.
[18] H. Qi, S. Xiao, R. Shi, M.P. Ward, Y. Chen, W. Tu, et al., COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis, Sci. Total Environ. (2020) 138778.
[19] R.C. Urban, L.Y.K. Nakada, GIS-based spatial modelling of COVID-19 death incidence in São Paulo, Brazil, Environ. Urbanization (2020), 0956247820963962.
[20] L.Y.K. Nakada, R.C. Urban, COVID-19 pandemic: environmental and social factors influencing the spread of SARS-CoV-2 in São Paulo, Brazil, Environ. Sci. Pollut. Res. (2020).
[21] IECDR, Bangladesh Covid-19 Update, Institute of Epidemiology, Disease Control and Research (IEDCR), 2020.
[22] NewAge. Bangladesh Confirms First Coronavirus Death Dhaka, 2020. Dhaka, Bangladesh [updated March 19].
[23] IMF, Bangladesh: Requests for Disbursement under the Rapid Credit Facility and Purchase under the Rapid Financing Instrument-Press Release; Staff Report; and Statement by the Executive Director for Bangladesh, International Monetary Fund, 2020.
[24] S. Anwar, M. Nasrullah, M.J. Hosen, COVID-19 and Bangladesh: challenges and how to address them, Front. Public Health 8 (154) (2020).
[25] WHO, COVID-19 Bangladesh Situation Reports, 2020.
[26] S.S. Preetha, Z. Islam, Is foul play the new normal? Daily Star (2020).
[27] S. Amit, Coronavirus and impact on the Bangladesh economy: assessing the damage of the black swan of 2020, Asia-Pacific J. 18 (15) (2020) 5441.
[28] LBAMCL, COVID-19 Impact on Bangladesh Economy, 2020. Dhaka.
[29] A. Mollalo, L. Mao, P. Rashidi, G.E. Glass, A GIS-based artificial neural network model for spatial distribution of tuberculosis across the continental United States, Int. J. Environ. Res. Publ. Health 16 (1) (2019) 157.
[30] J.B. Meigs, Invited commentary: insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor Analysis reveals patterns in the fabric of correlated metabolic risk factors, Am. J. Epidemiol. 152 (10) (2000) 908–911.
[31] R. Varraso, J. Garcia-Aymerich, F. Monier, N. Le Moual, J. De Batlle, G. Miranda, et al., Assessment of dietary patterns in nutritional epidemiology: principal component analysis compared with confirmatory factor analysis, Am. J. Clin. Nutr. 96 (5) (2012) 1079–1092.

[32] J. Merlo, B. Chaix, H. Ohlsson, A. Beckman, K. Johnell, P. Hjerpe, et al., A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena, J. Epidemiol. Community Health 60 (4) (2006) 290.
[33] M.J.G. Blyth, R. Kincaid, M.A.C. Craigen, G.C. Bennet, The changing epidemiology of acute and subacute haematogenous osteomyelitis in children, J. Bone Joint Surg. Br. 83-B (1) (2001) 99–102.
[34] A. Mollalo, A. Sadeghian, G.D. Israel, P. Rashidi, A. Sofizadeh, G.E. Glass, Machine learning approaches in GIS-based ecological modeling of the sand fly Phlebotomus papatasi, a vector of zoonotic cutaneous leishmaniasis in Golestan province, Iran, Acta Trop. 188 (2018) 187–194.
[35] D.A. Lovett, A.J. Poots, J.T.C. Clements, S.A. Green, E. Samarasundera, D. Bell, Using geographical information systems and cartograms as a health service quality improvement tool, Spatial Spatio-temporal Epidemiol. 10 (2014) 67–74.
[36] D.C. Wheeler, A. Páez, Geographically weighted regression, in: Handbook of Applied Spatial Analysis, Springer, 2010, pp. 461–486.
[37] T.M. Oshan, Z. Li, W. Kang, L.J. Wolf, A.S. Fotheringham, mgwr: a Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale, ISPRS Int. J. Geo-Inf. 8 (6) (2019) 269.
[38] M.D. Ward, K.S. Gleditsch, Spatial Regression Models, Sage Publications, 2018.
[39] C. Brunsdon, A.S. Fotheringham, M.E. Charlton, Geographically weighted regression: a method for exploring spatial nonstationarity, Geogr. Anal. 28 (4) (1996) 281–298.
[40] H. You, X. Wu, X. Guo, Distribution of COVID-19 morbidity rate in association with social and economic factors in Wuhan, China: implications for urban development, Int. J. Environ. Res. Publ. Health 17 (10) (2020).
[41] B. Roberts, T. Kanaley, Urbanization and Sustainability in Asia: Case Studies of Good Practice, Asian Development Bank, 2006.
[42] Wheaton WC, Kinsella Thompson A. The Geography of COVID-19 Growth in the US: Counties and Metropolitan Areas. Available at SSRN 3570540. 2020.
[43] A. Wilkinson, Local response in health emergencies: key considerations for addressing the COVID-19 pandemic in informal urban settlements, Environ. Urbanization 32 (2) (2020) 503–522.
[44] W.B. Bangladesh, Poverty Assessment: Assessing Progress from 2010 to 2016/17, 2019. Washington DC.
[45] A.M.N. Aziz, P. Atkins, Understanding the monga in northwest Bangladesh: household perceptions and perceptual connotations, Int. Res. J. Soc. Sci. 3 (8) (2014) 22–29.
[46] M. Rukunujjaman, Poverty level of Bangladesh: special case, Rangpur and Rajshahi division, Int. J. Soc. Res. Methodol. 5 (1) (2016) 212–228.
[47] M.A. Amin, Covid-19 job cuts forcing people to leave Dhaka, Dhaka Tribune (2020).
[48] S. Ali, N. Amin, Low-Income People Leaving Dhaka, The Business Standard, 2020.
[49] R. Ahamad, Families Leaving Dhaka as Coronavirus Crisis Hits Livelihood, New Age Bangladesh, 2020.
[50] F. Mahmud, Hundreds of doctors in Bangladesh infected with coronavirus, Aljazeera (2020).
[51] S. Mamun, Dhaka bracing for homebound rush for Eid amid pandemic, Dhaka Tribune (2020).
[52] S.N. Sakib, Police in Bangladesh at Great Risk from Pandemic, Anadolu Agency, 2020.