



Article

# Predicting COVID-19 Infections in Eswatini Using the Maximum Likelihood Estimation Method

Sabelo Nick Dlamini <sup>1,2,\*</sup> , Wisdom Mdumiseni Dlamini <sup>1</sup> and Ibrahima Socé Fall <sup>2</sup>

<sup>1</sup> Department of Geography, University of Eswatini, Kwaluseni, Manzini M200, Eswatini; wdlamini@uniswa.sz

<sup>2</sup> World Health Organization, 1211 Geneva, Switzerland; socef@who.int

\* Correspondence: sabelodlamini2004@yahoo.com

**Abstract:** COVID-19 country spikes have been reported at varying temporal scales as a result of differences in the disease-driving factors. Factors affecting case load and mortality rates have varied between countries and regions. We investigated the association between socio-economic, weather, demographic and health variables with the reported cases of COVID-19 in Eswatini using the maximum likelihood estimation method for count data. A generalized Poisson regression (GPR) model was fitted with the data comprising 15 covariates to predict COVID-19 risk in the whole of Eswatini. The results show that the variables that were key determinants in the spread of the disease were those that included the proportion of elderly above 55 years at 98% (95% CI: 97–99%) and the proportion of youth below the age of 35 years at 8% (95% CI: 1.7–38%) with a pseudo R-square of 0.72. However, in the early phase of the virus when cases were fewer, results from the Poisson regression showed that household size, household density and poverty index were associated with reported COVID-19 cases in the country. We then produced a disease-risk map of predicted COVID-19 in Eswatini using variables that were selected by the regression model at a 5% significance level. The map could be used by the country to plan and prioritize health interventions against COVID-19. The identified areas of high risk may be further investigated to find out the risk amplifiers and assess what could be done to prevent them.

**Keywords:** COVID-19; Eswatini; risk mapping; Poisson regression



**Citation:** Dlamini, S.N.; Dlamini, W.M.; Fall, I.S. Predicting COVID-19 Infections in Eswatini Using the Maximum Likelihood Estimation Method. *Int. J. Environ. Res. Public Health* **2022**, *19*, 9171. <https://doi.org/10.3390/ijerph19159171>

Academic Editors: Zhengchao Dong, Juan Manuel Gorriiz and Yudong Zhang

Received: 6 July 2022

Accepted: 20 July 2022

Published: 27 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

COVID-19 has spread dramatically since it was first discovered in China in December 2019 [1]. In as much as the spread resulted in a global pandemic [2], countries continued to report national spikes of infections at different temporal scales as the dominant strain of the virus took its toll on the population. Studies had reported various determinants of country COVID-19 spikes and they included factors such as population density, adherence to preventive measures and a host of typical socio-economic factors [3]. The disease spreads through direct and effective contact with infected persons; as a result, mechanical barriers limiting population interactions and mixing, such as social distancing, had been enforced by many countries [4]. Understanding the environmental risk factors, as well as the socio-economic and demographic factors associated with the spread of COVID-19, is crucial for effective prevention and response efforts [5]. Often, the underlying factors would not always be exactly the same between affected countries as population behavior, structure and socio-economic activities may vary from country to country [6]. For instance, the onset and severity of the disease and its associated country spike had often been seen by disproportionate case loads and mortality numbers between countries [7]. The factors affecting the spread of the pandemic, case load and mortality rates had been studied at national and subnational levels and the identified key drivers varied between countries and regions. A study on COVID-19 age mortality curves had shown that the spread of COVID-19 was different for high-income countries (HIMs) and low–middle-income

countries (LMICs) [8]. Interestingly, another study by [9] found that there were different susceptibilities and vulnerabilities to COVID-19 in urban and rural populations in the United States. These findings showed that the geographic context was a key determinant in the spread and severity of the disease.

Eswatini (formerly known by its English name as Swaziland), a country in the south-eastern Africa region with a population of about 1.1 million, had not been spared by the pandemic. According to the Ministry of Health, this lower-middle-income country had reported cumulative cases of about 69,000 and total deaths of about 1400 people by February 2022. The government of Eswatini had also implemented strict lockdowns and travel restrictions during the peaks of the pandemic in 2020 and 2021 [10]. However, cases continued to soar in the midst of these restrictions and, in this study, we were interested in determining the factors that were the potential drivers of infections during the peaks of the disease waves and surges in Eswatini. For instance, Eswatini is known to have high poverty proportions at 58.9%, where 20% is considered to be extremely poor, high HIV infections at 26% and high tuberculosis rates, which could have increased susceptibility to the disease because of co-morbidities and other demographic factors [11].

Previous COVID-19 studies on the disease predictors and case-load forecasting included a host of clinical, epidemiological, demographic, and socio-economic factors [12–14]. In this study, we explored the association between various socio-economic, weather, demographic and health variables with the distribution and spread of COVID-19 cases in Eswatini using the maximum likelihood estimation method for count data. Count data are dispersed data that involve discrete variables and require discrete analysis methods to estimate the parameters. We then mapped and predicted the risk of COVID-19 for the whole of Eswatini using those variables that were significantly associated with the disease. We believe that the mapping products produced will not only help in the ongoing health intervention efforts against COVID-19 but they will also help in the identification of the risk amplifiers of the disease in the affected areas. This work is a continuation from our previous work, which used a similar set of predictors to perform a spatial risk assessment and identified epidemic risk prone areas based on susceptibility risk, transmission risk and exposure risk [15].

## 2. Materials and Methods

### 2.1. Study Area and Data Sources

Eswatini is a southern African country bordered by South Africa all around except on the northeastern side where it is bordered by Mozambique. COVID-19 had spread throughout the entire country and its severity varied between the four districts comprised of Hhohho, Manzini, Shiselweni and Lubombo. Figure 1 shows the geographic location and distribution of reported COVID-19 cases in Eswatini while Figure 2 shows the incidence of cases per 1000 population. As can be seen in the maps, COVID-19 had spread throughout the rest of the country since the first case was introduced in March 2020, affecting both rural and urban populations, but its risk severity varied from locality to locality. In this study, we were interested in investigating the determinants of this risk variation by locality. The country is landlocked and travel within and outside Eswatini is mostly by road. Informal crossings (mainly to South Africa and Mozambique) are also common along the borders, especially for the local communities residing along the international boundaries.

To map and predict the distribution of COVID-19 in the country, we used data ranging from weather, communication, internet connectivity, traffic density, health and demographic, as well as socio-economic variables. The data were made up of a list of 15 covariates, which were used as predictors and regressed with georeferenced case data via a generalized Poisson regression (GPR) model [16]. Data variables on location entities such as supermarkets, shopping and church distances were extracted from Google Earth. The health, demographic and socio-economic data used in this study were obtained from the central statistics office (CSO). The CSO data included household density, household size, human immunodeficiency virus (HIV) prevalence, population density, youth proportion

and proportion of elderly above 55 years. Details for some of the data variables used in this study had been provided in our previous work [15]. COVID-19 case data comprising of reported cases ranging from the first case in March 2020 until December 2021 were obtained from the Ministry of Health [17]. The COVID-19 case data were made up of variables such as cases sex, age, case locality, clinical severity and date of hospital presentation. The data totaled 12,986 individual COVID-19 cases. The data were summed and aggregated at enumeration area (EA) level (made up of a total of 2326 EAs). The EAs serve as the smallest census units in the country, ranging from an area of about 0.013 km<sup>2</sup> to about 194 km<sup>2</sup>. The data used in this study are described and presented in Table 1.

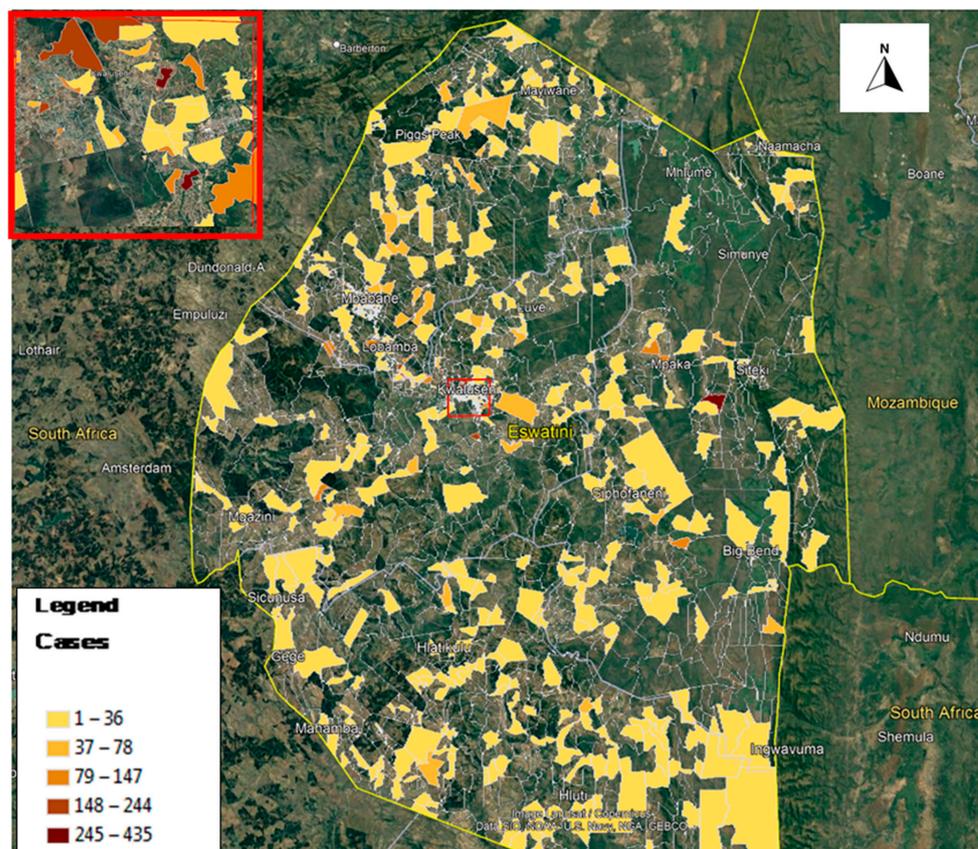


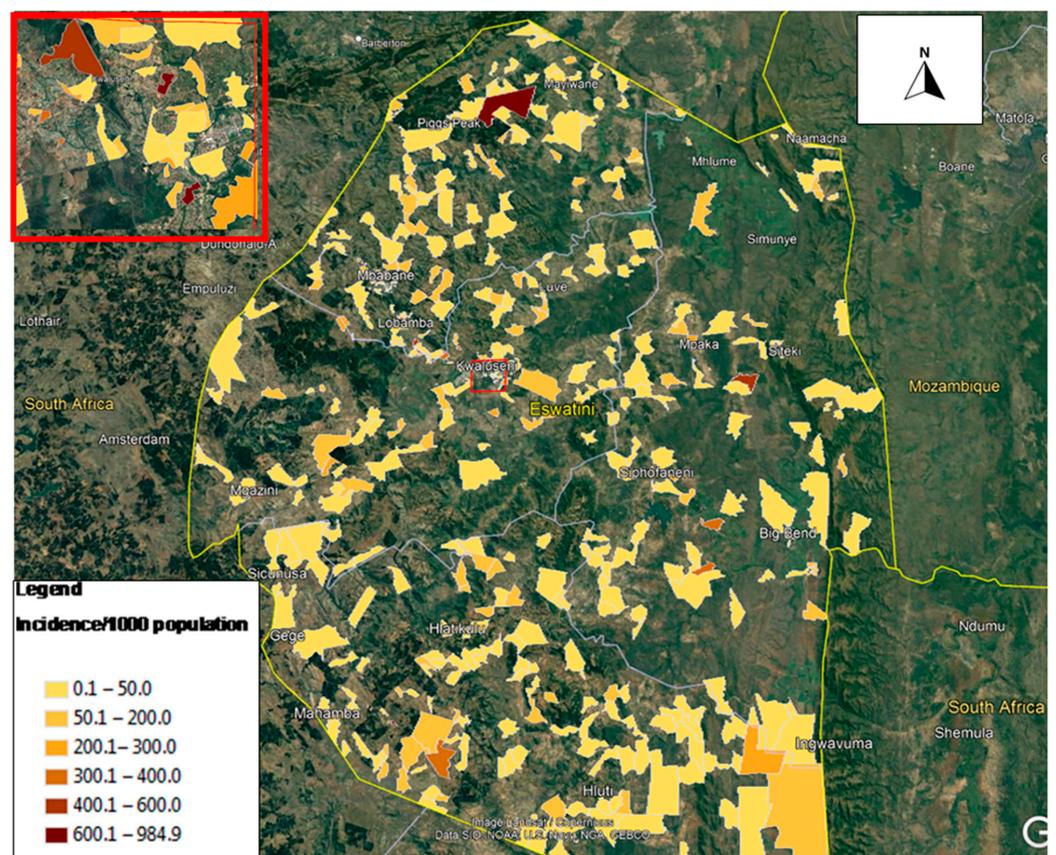
Figure 1. Distribution of reported COVID-19 cases in Eswatini.

Table 1. Health, Demographic and Socio-Economic Data.

No.	Variable	Variable Short Name	Format	Description
1.	Cellphone usage	Cellphone	Floating point Values ranging from 0 to 1	Proportion of cellphone users per EA
2.	Church-distance	Church_dis	Floating point Distance (km)	Distance between EA and church
3.	Elderly above 55 years	Elderly_55	Floating point Values ranging from 0 to 100	Percentage or number of people above 55 years of age per 1000 people in each EA
4.	Household density	Hhld_dens	Floating point Number of buildings per unit area (km <sup>2</sup> )	Numerical quantities of the built up surface area in each EA
5.	Household size	Hhld_size	Integer number Number of people per household	Number of persons living in a private dwelling unit
6.	HIV Prevalence	HIV_prev	Floating point Values ranging from 0 to 100	Percentage of people living with HIV in each EA

**Table 1.** Cont.

No.	Variable	Variable Short Name	Format	Description
7.	Internet connectivity	Internet	Floating point values ranging from 0 to 100	Percentage of people connected to internet either via a computer or other devices
8.	Poverty index	Po_index	Floating point Values ranging from 0 to 100	Percentage of people living below USD 2 per day in each EA.
9.	Building density	People_bld	Floating point Values ranging from 0 to 100	Percentage of built up area in each EA
10.	Youth proportion	Youth_prop	Floating point Values ranging from 0 to 100	Percentage or rate of people below 35 years per 1000 people of age in each EA.
11.	Shopping distance	Shop_dist	Floating point Distance (km)	Distance between EA and shopping area
12.	Supermarket distance	Supmkt_dis	Floating point Distance (km)	Distance between EA and supermarket
13.	Temperature	Temp	Floating point Degrees celcius	Hot/cold
14.	Traffic mean	Traff_mean	Floating point Number of vehicles moving through an area per day per unit area (km <sup>2</sup> )	Numerical quantities of average traffic moving through each EA approximated as a surface area of that EA
15.	Population density	Pop_dens	Floating point Number of people/per unit area (km <sup>2</sup> )	Numerical quantities of the populated surface area in each EA.
16.	Total population		Integer number Number of people	Number of people in the entire country obtained by summing up the number of people recorded in each EA



**Figure 2.** Distribution of COVID-19 cases per 1000 population in Eswatini.

## 2.2. Statistical Modeling

In the first stage of the model, we conducted a bivariate stepwise regression to select the set of variables that were associated with COVID-19 at a 5% significance level and then we fitted the selected variables into the GPR model to estimate the incidence rate ratio (IRR) of the virus in Eswatini. The stepwise regression method allows the user to evaluate variables according to their order of importance in explaining the outcome variable of interest [18]. This way, we were able to select useful or significant subsets of the regressed variables in addition to identifying any variable enhancement or suppression [19] by adding or removing different single variables or a combination of variables in the model. In the final stage we used the variables selected in the second stage of the model to predict COVID-19 cases in the entire country using the maximum likelihood estimation method for count data. Association of COVID-19 symptoms with age was also assessed by categorizing individual ages into different groups to assess the most affected patients by age. Before selecting the model, we first investigated which model would be suitable for fitting the data between two more-or-less similar models involving the Poisson model and the negative binomial model [20] by assessing the presence of over-dispersion in the data. We checked for excess variation in the model by assessing whether the deviance of the model was greater than its corresponding degree of freedom. We also wanted to find out if any excess variation was due to over-dispersion or if it was due to excess zeros that may emanate from zero-reported cases in some of the EAs [21]. Since the negative binomial model is suitable in cases where there is over-dispersion or excess zeros, in our case we proceeded with the Poisson model as the deviance was not greater than its degree of freedom. We then developed a Poisson regression model that was first fitted with all 15 covariates using STATA statistical software version 13 [22] to select the most parsimonious set of variables that were associated with COVID-19. Variables that were selected at a 5% significance level were then added into the GPR model to predict the risk of COVID-19 in the whole of Eswatini. The Poisson model has only one parameter,  $\mu$  (mean), which corresponds to the mean of the case counts. In our case, the counts (number of events) were weighted by the context in which they were occurring. In our study, we used the population in each EA as the denominator and the case counts that correspond to rates as the numerator, as shown below:

$$R = \frac{N}{P}, \text{ the Poisson regression then takes the form, } \log(R_i) = a + \beta X_i,$$

which is equivalent to:

$$\log(N_i) = \log(P_i) + a + \beta X_i \quad (1)$$

where  $(N_i)$  is the rate of count in each EA ( $i$ ) and  $P$  is the total population in that EA,  $a$  is the intercept and  $\beta X_i$  are the regression coefficients. In the GPR model, we used the maximum likelihood method to estimate the outcome of the counts of COVID-19 cases as follows:

Let  $\mu_{ij}$  be the mean count of  $Y_{ij}$  events, which in this model corresponds to rate of COVID-19 cases on each EA  $i$  on a particular date  $j$  and  $X_i \sim P(\mu_i, 1 - \mu_i)$  with likelihood:

$$P(Y_{ij}; \mu_{ij}) = \frac{\mu_{ij}^{Y_{ij}}}{Y_{ij}!} e^{-\mu_{ij}} \quad (2)$$

Since the counts for  $\mu$  can only be positive, we had to log transform the values so that the model can take both negative and positive values on either side of the equation. Therefore, when linearizing the model, we take the logarithm of  $\mu$  as follows:

$$\log(\mu_i) = \log(N_i) + a + \sum_{k=1}^{15} \beta X_k \quad (3)$$

where  $N_i$  is the total population in each EA used here as an offset to scale the modeling of rate  $\mu_i$  and  $\beta X$  are the regression coefficients corresponding to the 15 covariates. The Poisson regression, when applied on rates, assumes that the outcome is the count (numerator), while

the log of the denominator,  $\log(N_i)$ , is considered as a covariate with regression coefficient fixed to 1 [23]. The maximum likelihood estimation method determines the values for the parameters in the model by maximizing the likelihood of observing the reported number of cases in each EA. The parameter values under investigation are found by maximizing the likelihood that the process described by the model produced the data that were actually observed in the original dataset. The model then maximizes the probability of obtaining the data we observed by estimating the mean number of cases and assuming that the counts of cases in the EAs were independent of each other. In our model, we assumed that the mean counts of COVID-19 cases in space and time followed a Poisson distribution, which is a discrete frequency distribution that gives the probability of a number of independent events occurring in a fixed space and time. The second model involved the fitting of the data into the same Poisson model but with only those variables that were selected at a 5% significance level in the first model fitting. The selected variables were used to predict the risk of COVID-19 in the whole of Eswatini using ArcGIS mapping software. COVID-19 predictions conducted in STATA were joined by EA using the ArcGIS spatial join tool. We predicted COVID-19 in unsampled locations in the country using the kriging method found in the ArcGIS spatial analyst tool. The number of interpolation points was set to 3 to force the model to estimate from the nearest sampled location and the search radius was set at 100 m to identify locality clusters associated with COVID-19.

### 3. Results and Discussion

#### 3.1. Generalized Poisson Regression Model

The results from the Poisson modeling with all 15 predictors showed that two variables were associated with COVID-19 at a 5% significance level. This was partly due to the rest of the socio-economic variables not varying in space when the virus became indiscriminate in its attacks and spread throughout the entire country. Therefore, variation by these variables did not hold and they became insignificant, especially when most of the EAs had reported at least one case. The selected variables included the proportion of elderly above 55 years of age at 98% (95% confidence interval (CI): 97–99%) and the proportion of youth below 35 years at 8% (95% CI: 1.7–38%) with a pseudo R-square of 0.72 (Table 2). These results were consistent with the findings by [24] on the clinical manifestations of the disease by age where they found that older people (age > 60) were at greater risk of developing severe symptoms from COVID-19 compared to younger people. Age remained an important mortality risk factor in a study by [25], while another study by [26] found that elderly people aged over 60 years exhibited more severe symptoms and had higher mortality rates among patients infected with COVID-19. Indeed, the difference in the proportion of elderly compared to the youth may also be explained by the fact that a greater percentage of the elderly were likely to present with clinical symptoms at health facilities and get tested, compared to the youth who were found to be mostly asymptomatic. Furthermore, studies have found that COVID-19 infections can be driven by asymptomatic young people [27,28].

In this study, a bivariate analysis of the clinical symptoms by age of COVID-19 patients showed that an average of 33-year-old (95% CI: 20–46 years) patients had no symptoms (Table 3). This age group was part of the proportion of youth and the working group in Eswatini. It also represents the most active group in terms of mobility, especially for livelihood and employment pursuits. A study by [29] found that the median age of COVID-19 patients with mild or moderate symptoms was 28 years. The findings in this study proved that a greater proportion of the youth who were infected with COVID-19 did not have any symptoms and thereby were less likely to get tested as the government testing strategy tended to be more reactive, especially in the earlier phase of the pandemic when testing kits were embarrassingly scarce. Furthermore, in the early stages of the pandemic, medical attention was given to the screening and management of symptomatic patients, and asymptomatic patients were only attended much later as clinical understating of the transmission of the virus improved.

**Table 2.** Results of the Poisson model.

Individual	IRR	Std Err.	z	P > z	95% CI	
cellphone	3.336945	2.693032	1.49	0.135	0.6861192	16.229
church_dis	0.991159	0.016569	−0.53	0.595	0.9592107	1.0242
elderly_55	0.984678	0.0025786	−5.9	0.000 *	0.979637	0.9897
hhld_dens	1.000045	0.0001841	0.25	0.806	0.9996845	1.0004
hhld_size	0.9565653	0.0408021	−1.04	0.298	0.8798463	1.04
hiv_prev	0.3712601	0.3929613	−0.94	0.349	0.046636	2.9555
internet	0.936553	0.4262686	−0.14	0.885	0.3838055	2.2854
p0_index	1.00156	0.0044225	0.35	0.724	0.99293	1.0103
people_bld	0.9836968	0.0358074	−0.45	0.652	0.9159606	1.0564
pop_dens	1.000002	0.0000865	0.03	0.977	0.999833	1.0002
youth_prop	0.0816543	0.0646422	−3.16	0.002 *	0.0173029	0.3853
shop_dist	0.9976175	0.0178658	−0.13	0.894	0.9632086	1.0333
supmkt_dis	1.002583	0.0159068	0.16	0.871	0.9718864	1.0343
temp	0.9545037	0.0276899	−1.61	0.108	0.9017466	1.0103
traff_mean	0.9999643	0.0000665	−0.54	0.591	0.9998339	1.0001

\* Selected variables at 5% significance level; IRR = incident rate ratio, SD = standard deviation, z = Z score, P > z = significance-value or probability-value, 95% CI = 95% confidence interval for the estimated IRR.

**Table 3.** Mean age by symptoms.

Symptoms	Mean Age by Symptoms			
	Mean	SD	N	% N
No symptoms	33.4	13.75	5566	42.86
Mild	34.8	11.23	4681	36.05
Moderate to severe	48.0	16.97	177	1.36
Severe	29.0	12.35	88	0.68
Recovered	36.0	10.78	1236	9.52
Deceased	58.0	1.41	177	1.36
Unknown	36.2	9.39	1060	8.16
<b>Total</b>	<b>34.9</b>	<b>12.47</b>	<b>12,986</b>	<b>100</b>

Other variables such as church distance, shopping distance and supermarket distance were not significantly associated with COVID-19 infections and this may be attributed to the fact that most of these centers were closed and visiting such centers was severely restricted during the government's strict lockdown measures [30]. Surprisingly, household size, household density and poverty index were not associated with COVID-19. However, other studies have found that household size and population density were associated with COVID-19 [31–34]. We wish to clarify that in the initial phase of the virus when cases were fewer and emerging, preliminary results from the Poisson regression showed that these variables were associated with COVID-19. Similarly, studies using early COVID-19 data until December 2020 mostly had the same findings [35]. In Eswatini, this was mostly because it was highly likely to find a positive case within the household of an index case as the Ministry of Health was conducting contact tracing following a positive case. Moreover, cases were spreading faster among people living within low-cost housing and poorer but congested communities who had little or no room to self-isolate in the country and this finding was similar to what [36] found in Africa and Latin America. However, as infections had spread country-wide, affecting both rich and poor communities, the association of these variables with COVID-19 had been confounded. Moreover, the data did not vary in space as most of the EAs reported at least a single case. Also, the information, technology and communication variables such as internet connectivity and cellphone usage were also not associated with COVID-19. Some studies have shown that reduction in cellphone activity at work and retail locations was associated with lesser growth in COVID-19 cases [37]. Clearly, the spread of the virus was indiscriminate as it attacked the rest of the country.

### 3.2. Age and COVID-19 Infections

The results of the bivariate analysis on the distribution of symptoms by age revealed that younger people aged <35 years were less likely to present with symptoms. It is possible that a majority of infected young people did not get tested and may have recovered without being reported anywhere. Table 3 presents the results of the bivariate analysis between clinical symptoms severity and age of COVID-19 patients. Symptoms were categorized according to severity where: no symptoms referred to those who had a positive test but did not show any clinical manifestations of the disease; mild symptoms were those who had a positive test and were barely showing any symptoms and quickly returned to feeling normal; moderate to severe referred to individuals who had any of the virus signs and symptoms such as fever, cough and sore throat, among others; and severe symptoms referred to individuals who required emergency treatment and hospital admission. Severe symptoms included patients who reported with shortness of breath, feeling faint or passing out, and having severe chest pains. Although COVID-19 had severe clinical manifestations among an average of 29-year-old patients, it is worth mentioning that this was a very small percentage (0.68%). This finding could be partly attributed to the fact that the first wave of COVID-19 in Eswatini was spreading faster among the active youth who were working, in schools and in tertiary institutions in the country and often had severe clinical manifestations in some of them. There was also a small proportion (8.2%) of individuals whose symptoms were not immediately ascertained and they were categorized as “unknown” in the COVID-19 database.

Variables that were selected at 5% significance level were used in the GPR model to predict the risk of COVID-19 infections and Figure 3 shows the predicted risk of COVID-19 in the whole country. The risk of COVID-19 was found to be higher (180 cases per 1000 population) in the surrounding urban and peri-urban areas of main towns and industrial complexes of the country comprising Piggs Peak, Manzini, Mbabane, Nhlanguano, Bhunya, Mankayane, Matsapha, Siphofaneni and Big Bend. This was followed by rural and agricultural towns and other main places at 110 cases per 1000 population and these comprised Tshaneni, Simunye, Ngwenya and Sidvokodvo. The remainder of the rural areas and other sparse settlements in the countryside had less than 34 cases per 1000 population. A study by [38] also found that urban centers of the global south were highly likely to be predisposed to global risks such as COVID-19 because of their vulnerability and exposure, exacerbated by the process of urbanization in those centers. In Eswatini, urban centers are the backbone of daily activities and daily commuting between rural and urban centers is not uncommon. Indeed, cases reported in the earlier stages of the virus were mostly concentrated around urban and peri-urban areas and the daily contact between such places with rural commuters may have fueled the spread of the virus to the rest of the country. The remaining areas with lower or zero cases were mainly private farms and rangelands with low population density and no residential areas.

The investigation and understanding of the disease determinant factors by this study are not only crucial for the management and containment of COVID-19 but they are also crucial for the future management of similar pandemics. This present study investigated a collection of potential factors in the spreading of an infectious disease and found that there were disparities by age in its attack rate, symptom manifestations and mortality rate. Therefore, the social and demographic structure of a country determines the rate of the virus spread and its severity in the population. In addition, we found that it is important and also possible to manage infectious diseases such as COVID-19 in the very early stages of the onset of the virus when the driving factors could easily be predictable. For instance poverty, household density and household size may determine the rate of the disease propagation from an index case and studies have shown that quick isolation and quarantine of infected persons is crucial [39]. The spatial prediction in this study supports health intervention efforts as areas of heightened transmission can further be investigated to aid readiness and preparedness plan strategies even for similar future viruses. Our map products have shown that the virus has a potential to spread faster in the main places

where people live and interact, such as urban areas, compared to the rural areas that were characterized by sparse and dispersed settlements. The demographic factors investigated in this study, such as age, can also be used to plan, for instance, the ongoing vaccination program, whereby the right age category can be targeted and prioritized.

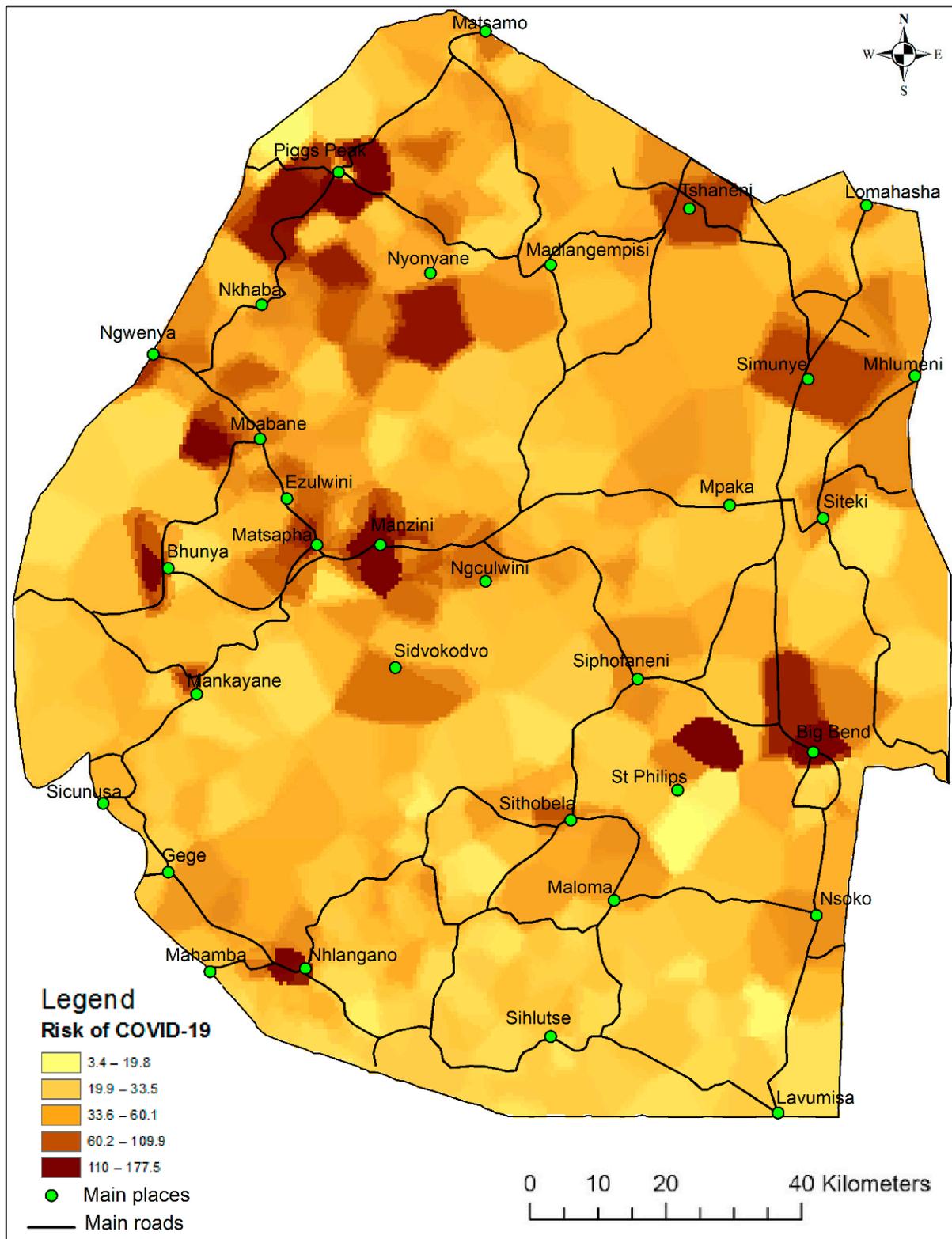


Figure 3. Predicted risk of COVID-19 infections in Eswatini.

#### 4. Conclusions

We implemented a maximum likelihood estimation method based on the GPR model to discover the set of socio-economic, weather, demographic and health variables that were associated with COVID-19 in Eswatini. We found that the proportion of elderly population above the age of 55 years and the proportion of youth population less than 35 years of age were highly associated with the disease. Clearly, the virus had severe symptom manifestations among the elderly, prompting them to seek medical attention and subsequently get tested for the disease. Therefore, the presence of a higher population of elderly was an indicator for risk to the virus as they were more likely to fall sick compared to the youth who were found to be mostly asymptomatic. In addition, we found that the higher the presence of youth within the population, the higher the spread of the virus even though the youth were mostly asymptomatic. Other future studies could investigate if the presence of asymptomatic people is not a precursor for the mutation of the virus, as we have seen with the evolving strains of the COVID-19 pandemic. The association of the disease with other socio-economic variables gets confounded as the virus indiscriminately spreads to the rest of the population in a country. While the first analysis of the disease conducted during the early stage of the pandemic when cases were fewer showed that it was associated with poverty, household size and household density; these associations did not hold as most of these variables were later not significant as cases spread. Caution must be taken when interpreting such results as they also depend on the stage of the virus and its level of spread in a country, and the confounding and interaction of variables must be investigated.

In this work, we mapped and predicted the risk of COVID-19 infections in Eswatini using socio-economic, weather, demographic and health variables. The mapping products produced in this work could be used by the country to plan and prioritize health interventions for similar diseases in the future. The areas of high risk may be further investigated to discover the risk amplifiers in those areas and to assess what could be done to prevent them. Our work contributes to the ongoing COVID-19 surveillance and response efforts in the country and to the rest of the global fight against the virus in areas with a similar setting. More research work is needed to investigate how some of the variables used in this study impact the effectiveness of control interventions and how they could be used to aid preparedness planning and readiness for future pandemics in the country and elsewhere.

**Author Contributions:** S.N.D. conceived the article, conducted formal analysis, investigation and methodology and wrote the manuscript. W.M.D. and I.S.F. conducted data analysis, and reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors declare that there was no funding received for this work.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Supporting data from this study could be requested directly from the corresponding author following a reasonable request.

**Conflicts of Interest:** The authors have no competing interest to declare that are relevant to the content of this article.

#### References

1. Chen, X.; Yu, B. First two months of the 2019 Coronavirus Disease (COVID-19) epidemic in China: Real-time surveillance and evaluation with a second derivative model. *Glob. Health Res. Policy* **2020**, *5*, 7. [[CrossRef](#)] [[PubMed](#)]
2. Khanna, R.C.; Cicinelli, M.V.; Gilbert, S.S.; Honavar, S.G.; Murthy, G.V.S. COVID-19 pandemic: Lessons learned and future directions. *Indian J. Ophthalmol.* **2020**, *68*, 703–710. [[CrossRef](#)]
3. Chaudhry, R.; Dranitsaris, G.; Mubashir, T.; Bartoszko, J.; Riazi, S. A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine* **2020**, *25*, 100464. [[CrossRef](#)] [[PubMed](#)]
4. Qian, M.; Jiang, J. COVID-19 and social distancing. *J. Public Health* **2022**, *30*, 259–261. [[CrossRef](#)] [[PubMed](#)]

5. Pradhan, D.; Biswasroy, P.; Kumar Naik, P.; Ghosh, G.; Rath, G. A Review of Current Interventions for COVID-19 Prevention. *Arch. Med. Res.* **2020**, *51*, 363–374. [[CrossRef](#)] [[PubMed](#)]
6. Hashim, M.J.; Alsuwaidi, A.R.; Khan, G. Population Risk Factors for COVID-19 Mortality in 93 Countries. *J. Epidemiol. Glob. Health* **2020**, *10*, 204–208. [[CrossRef](#)]
7. Sorci, G.; Faivre, B.; Morand, S. Explaining among-country variation in COVID-19 case fatality rate. *Sci. Rep.* **2020**, *10*, 18909. [[CrossRef](#)]
8. Demombynes, G. *COVID-19 Age-Mortality Curves Are Flatter in Developing Countries*; World Bank: Washington, DC, USA, 2020.
9. Peters, D.J. Community Susceptibility and Resiliency to COVID-19 Across the Rural-Urban Continuum in the United States. *J. Rural Health* **2020**, *36*, 446–456. [[CrossRef](#)]
10. Pasquali, G.; Godfrey, S. Governance of Eswatini Apparel Regional Value Chains and the Implications of Covid-19. *Eur. J. Dev. Res.* **2022**, *34*, 473–502. [[CrossRef](#)]
11. Nhapi, T.G.; Dhemba, J. The conundrum of old age and COVID-19 responses in Eswatini and Zimbabwe. *Int. Soc. Work* **2020**, *63*, 842–846. [[CrossRef](#)]
12. Gallo Marin, B.; Aghagoli, G.; Lavine, K.; Yang, L.; Siff, E.J.; Chiang, S.S.; Salazar-Mather, T.P.; Dumenco, L.; Savaria, M.C.; Aung, S.N.; et al. Predictors of COVID-19 severity: A literature review. *Rev. Med. Virol.* **2021**, *31*, e2146. [[CrossRef](#)] [[PubMed](#)]
13. Sun, Y.; Koh, V.; Marimuthu, K.; Ng, O.T.; Young, B.; Vasoo, S.; Chan, M.; Lee, V.J.M.; De, P.P.; Barkham, T.; et al. Epidemiological and Clinical Predictors of COVID-19. *Clin. Infect. Dis.* **2020**, *71*, 786–792. [[CrossRef](#)] [[PubMed](#)]
14. Moreira, A.; Chorath, K.; Rajasekaran, K.; Burmeister, F.; Ahmed, M.; Moreira, A. Demographic predictors of hospitalization and mortality in US children with COVID-19. *Eur. J. Pediatr.* **2021**, *180*, 1659–1663. [[CrossRef](#)]
15. Dlamini, W.M.; Dlamini, S.N.; Mabaso, S.D.; Simelane, S.P. Spatial risk assessment of an emerging pandemic under data scarcity: A case of COVID-19 in Eswatini. *Appl. Geogr.* **2020**, *125*, 102358. [[CrossRef](#)] [[PubMed](#)]
16. Consul, P.C.; Famoye, F. Generalized poisson regression model. *Commun. Stat.-Theory Methods* **1992**, *21*, 89–109. [[CrossRef](#)]
17. COVID-19 Eswatini. Available online: <http://datastudio.google.com/reporting/b847a713-0793-40ce-8196-e37d1cc9d720/page/2a0LB?feature=opengraph> (accessed on 6 July 2022).
18. Thompson, B. Stepwise Regression and Stepwise Discriminant Analysis Need Not Apply here: A Guidelines Editorial. *Educ. Psychol. Meas.* **1995**, *55*, 525–534. [[CrossRef](#)]
19. Currie, I.; Korabinski, A. Some Comments on Bivariate Regression. *J. R. Stat. Soc. Ser. Stat.* **1984**, *33*, 283–293. [[CrossRef](#)]
20. Gardner, W.; Mulvey, E.P.; Shaw, E.C. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychol. Bull.* **1995**, *118*, 392–404. [[CrossRef](#)]
21. Greene, W.H. *Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models*; Social Science Research Network: Rochester, NY, USA, 1994.
22. StataCorp. *Stata Statistical Software: Release 13*; StataCorp LP: College Station, TX, USA, 2013.
23. Schober, P.; Vetter, T.R. Count Data in Medical Research: Poisson Regression and Negative Binomial Regression. *Anesth. Analg.* **2021**, *132*, 1378–1379. [[CrossRef](#)]
24. Farshbafnadi, M.; Kamali Zonouzi, S.; Sabahi, M.; Dolatshahi, M.; Aarabi, M.H. Aging & COVID-19 susceptibility, disease severity, and clinical outcomes: The role of entangled risk factors. *Exp. Gerontol.* **2021**, *154*, 111507. [[CrossRef](#)]
25. Leung, C. Risk factors for predicting mortality in elderly patients with COVID-19: A review of clinical data in China. *Mech. Ageing Dev.* **2020**, *188*, 111255. [[CrossRef](#)] [[PubMed](#)]
26. Wei, C.; Liu, Y.; Liu, Y.; Zhang, K.; Su, D.; Zhong, M.; Meng, X. Clinical characteristics and manifestations in older patients with COVID-19. *BMC Geriatr.* **2020**, *20*, 395. [[CrossRef](#)] [[PubMed](#)]
27. Han, D.; Li, R.; Han, Y.; Zhang, R.; Li, J. COVID-19: Insight into the asymptomatic SARS-CoV-2 infection and transmission. *Int. J. Biol. Sci.* **2020**, *16*, 2803–2811. [[CrossRef](#)]
28. Balasubramanian, S.; Rao, N.M.; Goenka, A.; Roderick, M.; Ramanan, A.V. Coronavirus Disease 2019 (COVID-19) in Children—What We Know So Far and What We Do Not. *Indian Pediatr.* **2020**, *57*, 435–442. [[CrossRef](#)] [[PubMed](#)]
29. Lu, J.; Peng, J.; Xiong, Q.; Liu, Z.; Lin, H.; Tan, X.; Kang, M.; Yuan, R.; Zeng, L.; Zhou, P.; et al. Clinical, immunological and virological characterization of COVID-19 patients that test re-positive for SARS-CoV-2 by RT-PCR. *eBioMedicine* **2020**, *59*, 102960. [[CrossRef](#)]
30. Pillay, J. COVID-19 Shows the Need to Make Church More Flexible. *Transformation* **2020**, *37*, 266–275. [[CrossRef](#)]
31. Sigler, T.; Mahmuda, S.; Kimpton, A.; Loginova, J.; Wohland, P.; Charles-Edwards, E.; Corcoran, J. The socio-spatial determinants of COVID-19 diffusion: The impact of globalisation, settlement characteristics and population. *Glob. Health* **2021**, *17*, 56. [[CrossRef](#)]
32. Liu, P.; McQuarrie, L.; Song, Y.; Colijn, C. Modelling the impact of household size distribution on the transmission dynamics of COVID-19. *J. R. Soc. Interface* **2021**, *18*, 20210036. [[CrossRef](#)]
33. Hamidi, S.; Hamidi, I. Subway Ridership, Crowding, or Population Density: Determinants of COVID-19 Infection Rates in New York City. *Am. J. Prev. Med.* **2021**, *60*, 614–620. [[CrossRef](#)]
34. Sugg, M.M.; Spaulding, T.J.; Lane, S.J.; Runkle, J.D.; Harden, S.R.; Hege, A.; Iyer, L.S. Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Sci. Total Environ.* **2021**, *752*, 141946. [[CrossRef](#)]
35. Gershengorn, H.B.; Patel, S.; Shukla, B.; Warde, P.R.; Bhatia, M.; Parekh, D.; Ferreira, T. Association of Race and Ethnicity with COVID-19 Test Positivity and Hospitalization Is Mediated by Socioeconomic Factors. *Ann. Am. Thorac. Soc.* **2021**, *18*, 1326–1334. [[CrossRef](#)] [[PubMed](#)]

36. Bargain, O.; Aminjonov, U. Poverty and COVID-19 in Africa and Latin America. *World Dev.* **2021**, *142*, 105422. [[CrossRef](#)] [[PubMed](#)]
37. Sehra, S.T.; George, M.; Wiebe, D.J.; Fundin, S.; Baker, J.F. Cell Phone Activity in Categories of Places and Associations With Growth in Cases of COVID-19 in the US. *JAMA Intern. Med.* **2020**, *180*, 1614–1620. [[CrossRef](#)]
38. Shekhar, H.; Rautela, M.; Maqsood, M.; Paris, R.; Flores de León, R.M.; Romero-Aguirre, M.F.; Balinos, M.; Velázquez, M.E.; Amri, G.S.; Rahman, T.; et al. Are leading urban centers predisposed to global risks- An analysis of the global south from COVID-19 perspective. *Habitat Int.* **2022**, *121*, 102517. [[CrossRef](#)]
39. Chung, S.-C.; Marlow, S.; Tobias, N.; Alogna, A.; Alogna, I.; You, S.-L.; Khunti, K.; McKee, M.; Michie, S.; Pillay, D. Lessons from countries implementing find, test, trace, isolation and support policies in the rapid response of the COVID-19 pandemic: A systematic review. *BMJ Open* **2021**, *11*, e047832. [[CrossRef](#)] [[PubMed](#)]