Tanzania Journal of Science 47(3): 1266-1281, 2021
ISSN 0856-1761, e-ISSN 2507-7961

Spatial Modelling for Cholera Incidence Rates in Morogoro Municipality, Tanzania

Caroline Ngereza[1]*, Rosalia S. Katapa and Ali R. Mniachi
[1]*Department of Population Studies, Institute of Rural Development Planning, Dodoma, Tanzania*
[2]*Department of Statistics, University of Dar es Salaam, Dar es Salaam, Tanzania*
E-mail addresses: cngereza@irdp.ac.tz, rskatapa@yahoo.com, amniachi@gmail.com
**Corresponding author*
Received 25 May 2021, Revised 16 Aug 2021, Accepted 18 Aug 2021, Published Aug 2021
DOI: https://dx.doi.org/10.4314/tjs.v47i3.34

Abstract

Spatial modelling was conducted to examine community factors associated with cholera incidence rates in Morogoro Municipality. The study employed both secondary (cholera cases) and primary (geographic coordinates of community risk factors) data. Spatial lag model was applied in examining association between the variables. All wards of Morogoro Municipality were considered in the study to capture their variations because cholera cases have a tendency to be clustered. Results indicated that market density, distance to the market and distance to the dumpster are significant factors associated with cholera incidence rates in the wards ($p < 0.05$). Geographically weighted Poisson model was used to show the variations of those factors between the wards in Municipality. A statistically significant positive association of cholera incidence rates; and market density was only found in Mazimbu ward ($p < 0.05$) and distance from the community to the dumpster was found in Kihonda, Kingolwira, Bigwa, Kichangani, Kilakala and Boma wards ($p < 0.001$) and some wards at the centre of the municipality which are Mji Mkuu and Kingo ($p < 0.05$). A statistically significant negative association of cholera incidence rates and distance from the centre of the community to the market was found in Kihonda, Kingolwira and Kichangani ($p < 0.001$) and Bigwa wards ($p < 0.05$). Therefore, measures taken to control and prevent cholera disease should base on the variations of the risk factors found in the Municipal wards.

Keywords: Cholera incidence rate, Spatial lag regression model, Community risk factor, Geographically weighted Poisson model.

Introduction

Cholera remains a global problem to public health and a sign of imbalance and lack of social development. Worldwide, it is estimated that 2.9 million cases and 95,000 deaths due to cholera occur yearly (Ali et al. 2015). For a long time, the Asian sub-continent was the home of cholera (Harris et al. 2012). Currently, there is persistence of cholera outbreak in Africa with a large disease burden and high rates of case fatality. The majority of the reported cases of cholera are from Sub-Saharan Africa which contributes about 60% of cases and deaths globally (Ali et al. 2015). Since the beginning of the seventh pandemic in 1961, cholera does not affect African countries equally (Mengel et al. 2014, Clemens et al. 2017). From 1970 to 2011, half of all cholera cases from African countries were only from seven countries; Angola (183,076), Democratic Republic of the Congo (391,524), Mozambique (315,295), Nigeria (260,966), Somalia (255,788), Tanzania (204,569), and South Africa (186,462) (Mengel et al. 2014). In

Tanzania, the first cholera incidence was known or reported in 1974, and since then it has been reported almost every year with an approximation of 5,800 cases (URT 2014). From 1974 to 2018, Tanzania had reported over 250,000 cholera cases and 13,078 deaths (Hounmanou et al. 2019).

Several studies have been done on cholera disease globally. For example, Trærup et al. (2011) using Poisson regression model to find an association between the increase in temperature and the cholera incidence in Tanzania, found a significant relationship between temperature and the incidence of cholera. For a 1 degree Celsius temperature increase, the initial relative risk of cholera increases by 15 to 29 percent. Cowman et al. (2017) did multivariate analysis to investigate the relationship between the occurrence of cholera and various risk factors in Kenya. They found that, the risk of cholera was associated with open defecation, use of unimproved water sources, poverty headcount ratio and the number of health facilities per 100,000 populations. Identification of risk factors for the cholera outbreak in Yemen using logistic regression analysis was done by Dureab et al. (2019). The study found that, not washing khat and the use of common-source water in the household were significant risk factors for having cholera cases. In Nigeria, a study done by Olanrewaju and Adepoju (2017) investigated the spatial relationship between cholera incidences and environmental risk factors which found that waste dump sites and markets were highly associated with cholera incidences. Surface runoff from sources (waste dump sites, pit-latrines, and wastewater) may also cause an increase in contamination of water sources, while slow flowing and stagnation of water ways may lead to presence of pathogenic bacteria which cause diarrhoea diseases. There are various influential factors of cholera which lead to current trends of cholera emergence and re-emergence. Wahed et al. (2013) showed that, the disease occurs in epidemics when conditions of poor sanitation, crowding and famine are present, and it is common where minimum requirements of clean water and sanitation are not met.

In Tanzania, despite the increasing number of households using improved water sources from 54% in 2010/2011 to 61% in 2015/2016, and decreasing number of households using unimproved toilet facilities from 76% in 2011/12 to 65% in 2015/2016 (URT and ICF 2016); yet there was a re-emergence of cholera outbreak at the end of 2015 to 2018 in almost all regions of Tanzania mainland. Country-wise, although the numbers of cholera cases and deaths have declined, the current trends of emergence and re-emergence of cholera periodically are alarming, especially in some regions such as Morogoro, Iringa, Kigoma, Dodoma, Arusha, Rukwa, Manyara, Songwe and Ruvuma (UNICEF 2018). This implies that, cholera remains a threat in Tanzania and interventions at household level have failed to overcome the problem. Therefore, there is a need of examining risk factors associated with cholera incidence rates at community levels as well. Statistical spatial model is used in this study, since cholera cases have a tendency to be clustered in specific areas and among specific population groups (You et al. 2013).

Materials and Methods

Study area

The study was conducted in Morogoro municipality, Morogoro region. The region was selected since is one of the regions which were among cholera hotspot regions in the country (Hounmanou et al. 2019) and experienced emergence and re-emergence of cholera periodically (UNICEF 2018). Morogoro municipality is the capital city of Morogoro region. The region is divided in 9 districts; Morogoro municipality, Ulanga, Kilosa, Gairo, Ifakara, Kilombero, Mvomero, Malinyi and Morogoro Rural. Morogoro municipality is the only one with urbanisation status in the region (URT 2015). It is located between longitudes 37˚34'52"E and 37˚45'25"E and between latitudes 6˚38'56"S and 6˚55'8"S in Morogoro region. Morogoro municipality constitutes 19 administrative wards. On the north and west is bounded by Mvomero District and to the east and south is bounded by Morogoro Rural District (Figure 1). Nearly two thirds (65%) of the population in the Municipality live in un-serviced and unplanned settlements (Unhabitat 2009), which leads to an increase of waterborne diseases (Steven and Mkonda 2015). Most of the wastes produced in the Municipality are not collected; are either buried, disposed at roadside dumpsite or other dumping methods, and only 35% are collected and dumped (Mollel 2016).

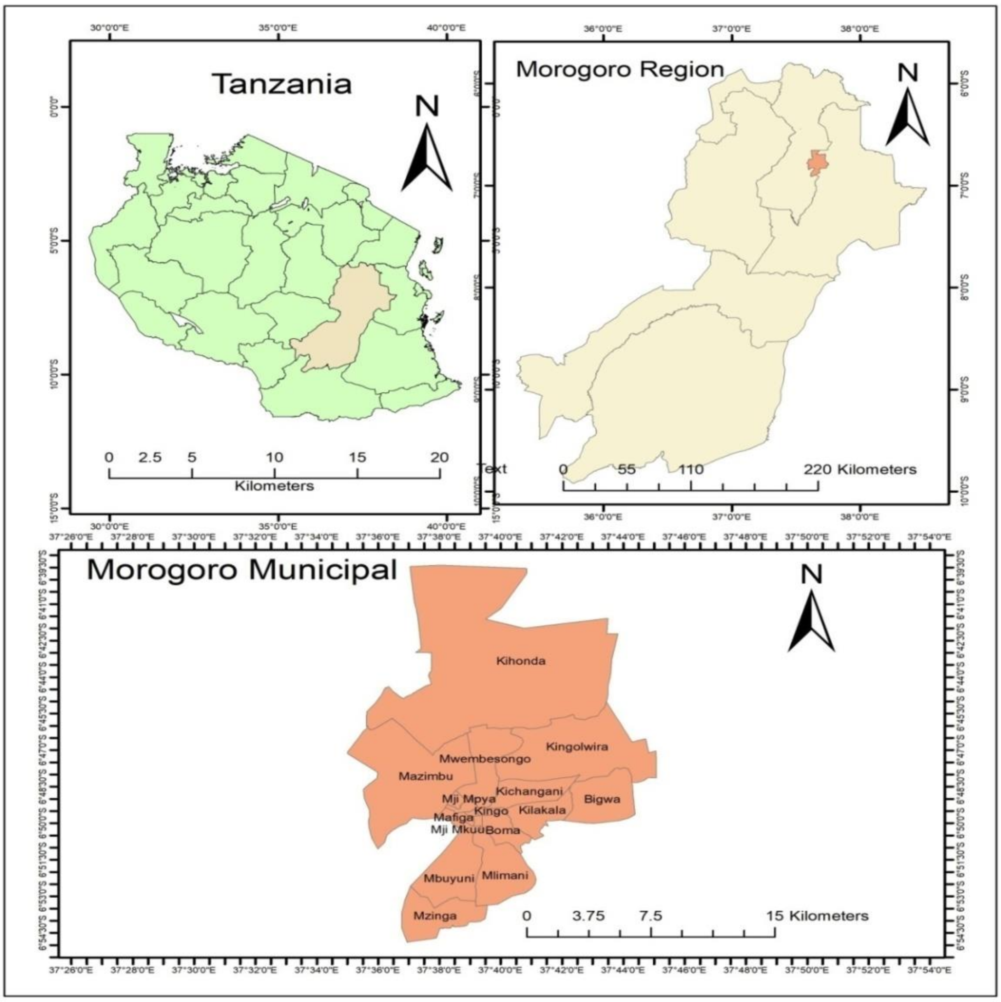

Figure 1: Map of Morogoro Municipal (Source: NBS 2012).

Research design and data

This study applied retrospective-cross sectional designs to examine possible associations between cholera incidence rates and community risk factors. The study used secondary and primary data. Secondary data were obtained from Morogoro Municipality Council, whereby cholera cases of all wards of Morogoro Municipal from 2015 to 2018 were collected. Primary data were collected by Handheld Global Positioning System (GPS) Garmin 12XL in order to determine the geographic coordinates (latitudes and longitudes) of all community risk factors (markets, slaughterhouses, dumpsters and dumpsites).

Study variables

Cholera incidence rate were the dependent variable to be measured. Incidence rate of each ward was calculated as the number of cholera cases in each ward (Kihonda, Kingolwira, Mazimbu, Mji Mkuu, Mji Mwema, Kilakala, Mwembesongo, Kiwanja cha Ndege, Uwanja wa Taifa, Kingo, Mafiga, Sultan Area, Boma, Mji Mpya, Kichangani, Bigwa, Mbuyuni, Mlimani, Mzinga) divided by averaged projected population of 2015 to 2018. In calculating disease rates for many infectious diseases, the numerator is the number of disease cases, while the denominator is the population of a given study area (Baig 2017). Cholera incidence raw rate was then multiplied by 10,000; which stands as cholera incidence rate per 10,000 people. Explanatory variables included were slaughterhouse density, market density, dumpsite density, dumpster density, distance to slaughterhouse, distance to market, distance to dumpsite, and distance to dumpster.

Data analysis

Data were analysed by using ArcGIS and R softwares. In examining spatial relationship between cholera incidence rates and community risk factors, spatial variables were first generated; by calculating Euclidean distance which showed the distance from the centre of community to each nearby community risk factor. Kernel density analysis also was done in calculating the density of each community risk factor, the area density unity used was square kilometers. Second, shape files containing cholera incidence rates, density variables and distance variables were merged to create one shape file which was used for regression analysis. Moran's I statistics were calculated for testing of spatial dependency of error term on OLS model. Lagrange multiplier test also was done in the selection of appropriate spatial model to apply (spatial lag model or spatial error model). Thereafter, geographically weighted Poisson regression model were generated based on significant variables of spatial lag model. Before generating the spatial regression model, all explanatory variables were also tested for multicollinearity.

The variance inflation factor used for testing presence of multicollinearity for the explanatory variable X_j is given by;

$$VIF_j = \frac{1}{1 - R_j^2}$$

where: R_j^2 is the coefficient of determination of variable X_j. When $VIF = 1$ it means no correlation, $1 < VIF \le 5$ means moderate correlation and $VIF > 5$ means high correlation, meaning that its effects should be controlled (Daoud 2017).

The kernel density function in this study was based on the quadratic kernel function (Silverman 1986), which is expressed as;

$$Density = \frac{1}{b^2} \sum_{i=1}^{n} \left[\frac{3p_i}{\pi} \left(1 - \left(\frac{d}{b} \right)^2 \right)^2 \right]$$

where: i are the input points (risk factors) which included only within the radius distance of the (u,v) ward location, p_i is the population of point i (optional parameter), d is the distance between point i and the (u,v) ward location and b is the bandwidth.

Spatial dependence

The spatial dependence is a functional relationship of what happens at one spatial point and what happens in another place (Anselin 1988). The Moran's I in a regression context is mostly used in diagnostic test. For the spatial model selection between spatial lag model and spatial error model, Langrage Multiplier (LM) test is used based on Anselin (1988) guidance. When Moran's I statistic for the error term of Ordinary Least Square **(**OLS) model is significant, it implies that there is spatial dependency in the model. After that, LM test for spatial lag model and spatial error model is done; when one test is significant, respective model is the best fit model. When both tests are significant, further test is conducted using Robust Langrage Multiplier (RLM), and the one with the lowest p-value is more applicable for the data.

Moran's I test is calculated as $I = (\frac{N}{SO})(\frac{\underline{e}'W\underline{e}}{\underline{e}\underline{e}'})$

where: $\underline{e}$ is an N by 1 vector of regression residuals from OLS estimation on N observations and W is N by N weight matrix, SO is the aggregate of spatial weight, i.e., $SO = \sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}$. Under Moran's I test of OLS error term, the hypothesis tested is; there is no spatial dependency between error terms. When the p-value is small ($p < 0.05$); the null hypothesis is rejected in favour of alternative hypothesis (there are spatial dependencies).

LM lag statistic is calculated as

$$LM(lag) = \frac{\left(\frac{e'Wy}{\delta^2}\right)^2}{\left((WX\beta)'\frac{MWX\beta}{\delta^2}\right) + tr\left(W'W + W^2\right)}$$

LM error statistic is given by

$$LM(error) = \frac{\left(\frac{\underline{e}'W\underline{e}}{\delta^2}\right)^2}{tr\left(W'W + W^2\right)}$$

where: y represent dependent variable, M = I - X(X'X)-1X', β is a vector with OLS estimates for the regression coefficients, W is N by N weight matrix and tr stands for the matrix trace operation (sum of diagonal elements), $\underline{e}$ is an N by 1 vector of a regression residuals from OLS estimation on N observations and W is N by N weight matrix. These statistics are asymptotically distributed as χ^2 with one degree of freedom.

Euclidean distance is calculated as;

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}$$

u_i and v_i are the geographic coordinates (longitude and latitude) in ward i and u_j and v_j are the geographic coordinates (longitude and latitude) in ward j.

Spatial Lag Regression Model (SLRM)

Spatial lag models are the inclusion of spatial lag variables to explain spatial dependence caused by externalities and spillover effects. The term "lag" refers to variables in target spatial data or a given space, which influence the spatial data of neighbouring areas. The spatial data of neighbouring areas also influence the target spatial data (Anselin 1988). The SLRM is expressed as;

$$y_i = \alpha + \rho\sum_{j=1}^{N} W_{ij} y_j + X_i\beta + \varepsilon_i$$

where: i is the spatial unit, y_i is the dependent variable (cholera incidence rate) at ward i; α is the intercept term, ρ is the spatial autoregressive coefficient of lag variable $\sum_{j=1}^{N} W_{ij} y_j$, ρ value is [0,1], and W_{ij} is a spatial weights matrix corresponding to ward i and j, and $\sum_{j=1}^{N} W_{ij} = 1$ for all value of i, X_i is the explanatory variable (distance variables and density variables) at ward i; β is the spatial regression coefficient and ε_i is the error term. The SLRM eliminates the intervention caused by spatial autocorrelation and tests the effects of spatial interactions. Parameter of spatial lag model is estimated by Maximum likelihood (ML) method (Anselin 1988). The log likelihood function is given by

$$\ln L = -\frac{N}{2}\ln\left(\pi\delta^2\right) + \ln\left|\mathrm{I} - \rho W\right| - \frac{e'e}{2\delta^2};$$

since $e = (\mathrm{I} - \rho W)y - X\beta$

where: I is N by N identity matrix, ML estimate for ρ is obtained from numerical optimization of a concentrated log-likelihood function.

Spatial Error Regression Model (SERM)

Spatial error model incorporates the presence of spatial autocorrelation in the model. It is expressed as:

$$y_i = \alpha + \lambda \sum_{j=1}^{N} W_{ij}\mu_j + X_i\beta + \varepsilon_i$$

where: i is the spatial unit, y_i is the dependent variable (cholera incidence rate) at ward i; α is the intercept term, λ is the spatial error coefficient, W_{ij} is a spatial weights matrix corresponding to ward i,j, μ_j is the error vector of a random error term assumed to be i.i.d, X_i is the explanatory variable (distance variables and density variables) at ward i; β is the spatial regression coefficient and ε_i is the modified error term. Parameters of spatial error model were also estimated by ML method. The log likelihood function is expressed as:

$$\ln L = -\frac{N}{2}\ln\left(\pi\delta^2\right) + \ln\left|\mathrm{I} - \lambda W\right|' - \frac{e'e}{2\delta^2};$$

$$\text{since } e = \left(\mathrm{I} - \lambda W\right)\left(Y - X\beta\right)$$

where: I is N by N identity matrix, ML estimate also is obtained from numerical optimization of a concentrated log-likelihood function.

Geographically Weighted Poisson Regression (GWPR)

The GWPR model was applied to show the local variation of the significant variables of spatial global model (spatial lag model). The model takes the form of:

$$y_i \sim p\left(e^{\sum_{k=1}^{p} \beta_k (u_i, v_i) x_{ik}} \right)$$

where; y_i is the cholera incidence rate of ward i, (u_i, v_i) are the coordinates (latitude and longitude) of ward i, x_{ik} is the k^{th} explanatory variable of region i, β_k is the coefficient of the k^{th} explanatory variable in region i.

A decision process was conducted to determine the most appropriate model (SLRM or SERM) to estimate the spatial data modified from (Anselin 1988). Geographically weighted regression model was added to the framework to capture the variations of dependent variables and explanatory variables across the wards as illustrate in Figure 2.

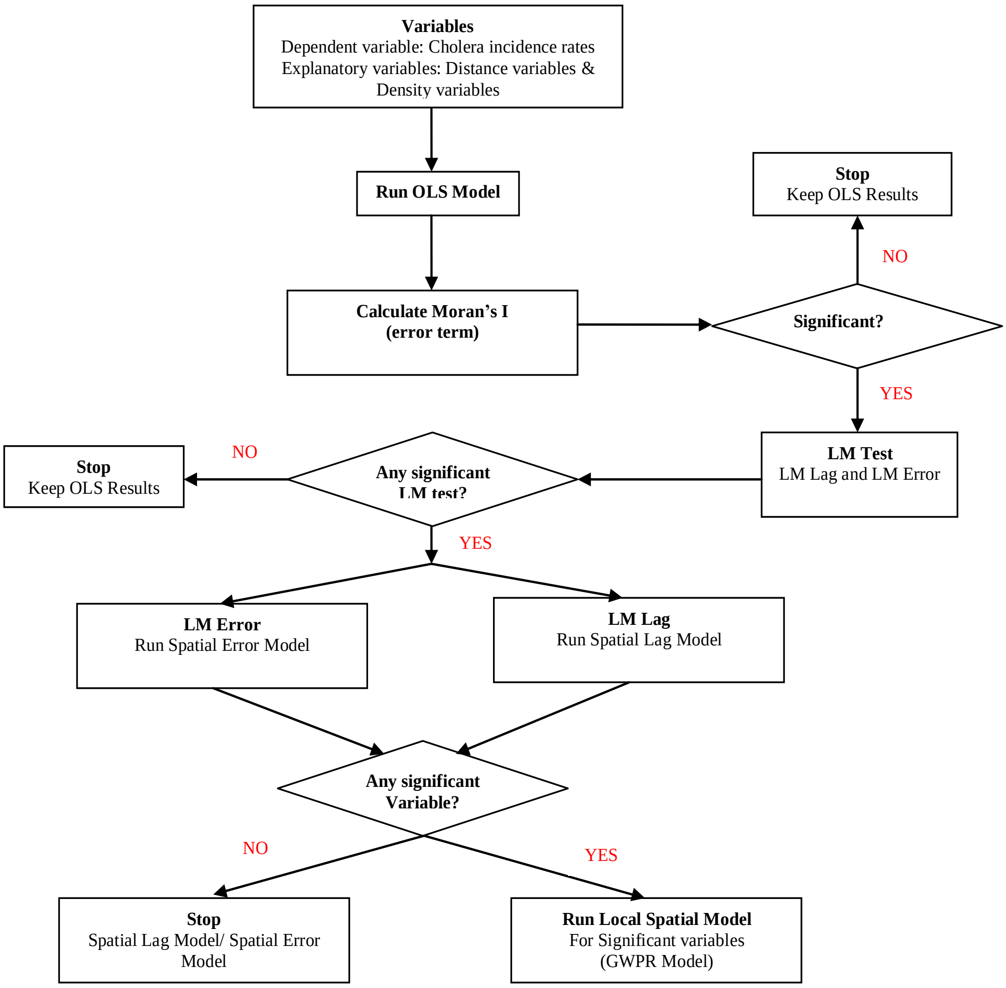

Figure 2: Decision process for spatial regression model (Source: Modified from Anselin 1988).

Results and Analysis

The summary statistics of the study variables are shown in Table 1. Cholera incidence rates ranged from 0 to 31.39 per 10,000 people with average of 5.34 and standard deviation of 7.56. The minimum value for each variable was zero, and wards with high cholera incidence rates were Kihonda and Kingolwira (Figure 3). The average densities of slaughterhouses, markets, dumpsters and dumpsites were 0.03, 0.36, 3.92 and 7.00, respectively. Average Euclidean distance from the centre of community to slaughterhouses, markets, dumpsters and dumpsites were 0.001, 0.01, 0.01 and 0.0001, respectively.

Table 1: Summary statistics of variables

Variable	Mean	Standard deviation	Maximum
Incidence rate per 10,000 people	5.34	7.56	31.39
Slaughterhouse density (Slaughterhouse per km^2)	0.03	0.11	0.43
Market density (Market per km^2)	0.36	0.40	1.45
Dumpster density (Dumpster per km^2)	3.92	4.71	14.83
Dumpsite density (Dumpsite per km^2)	7.00	30.51	133.00
Distance to slaughterhouse (km)	0.001	0.002	0.01
Distance to market (km)	0.01	0.01	0.04
Distance to dumpster (km)	0.01	0.01	0.03
Distance to dumpsite (km)	0.001	0.003	0.01

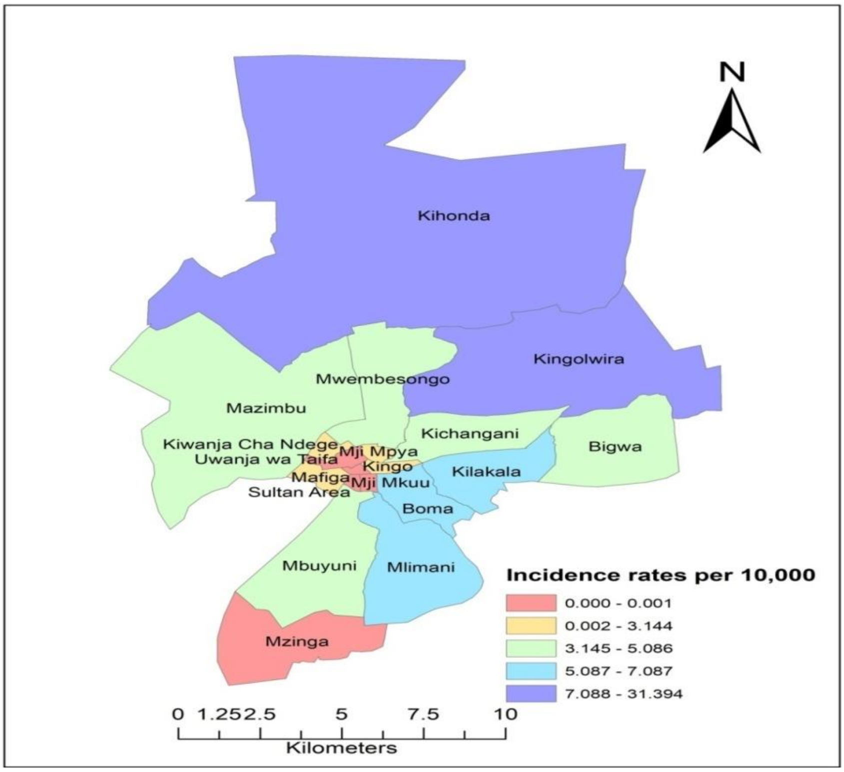

Figure 3: Graduated colour map showing cholera incidence rates for each ward.

Spatial dependence diagnostic test

Table 2 shows spatial dependence diagnostic test of the error term of OLS model according to the flow diagram shown in Figure 2. For density variables, Moran's I statistic had positive value of 0.186 with a $p<0.001$. This result indicates that spatial dependency and positive spatial autocorrelation exists in the model, and consequently, biased and inconsistent estimators of OLS model. The LM test was thereafter conducted to detect whether the spatial dependencies are in the lag, error or both (lag and error). The results revealed that, LM Lag p-value: $0.001 < 0.05$ and LM Error p-value: $0.068 > 0.05$. For distance variables, Moran's I statistic also had a positive value of

0.481 with a p=0.001. Also, LM Lag p-value: 0.004 < 0.05 and LM Error p-value: 0.444 > 0.05 was detected. Therefore, SLRM was recommended over SERM.

Table 2: Spatial dependence diagnostic test

	Kernel density variables		Average distance variables	
Tests	Value	P- value	Value	P- value
Moran's I	0.186	<0.001	0.481	0.001
Lagrange Multiplier (lag)	8.316	<0.001	8.0143	0.004
Lagrange Multiplier (error)	3.315	0.068	0.585	0.444

Table 3 shows that all the variables have VIF values less than 5 and corresponding tolerance statistic is more than 0.2. This implying that there is moderate association among the explanatory variables. Thus, all variables are suitable in spatial model fitting.

Table 3: Multicollinearity detection on variables

	Kernel density variables		Average distance variables	
Variable	VIF	Tolerance	VIF	Tolerance
Slaughterhouse	1.51	0.66	1.09	0.92
Market	1.51	0.66	4.05	0.25
Dumpster	2.73	0.37	3.55	0.28
Dumpsite	1.63	0.61	1.23	0.81

Results in Table 4 indicate that market density, distance to the market and distance to the dumpster are significant factors associated with cholera incidence rates since their corresponding p-values are less than 0.05. Also, the positive coefficients of market density and distance to the dumpster show that as a particular factor increases so does the cholera incidence rates in the ward increases. A negative coefficient of distance to the market shows that, as distance to the market increases, cholera incidence rates in the ward decreases.

Table 4: Spatial lag analyses of community risk factors related to cholera incidence rates

Parameter	Kernel density variables				Average distance variables			
	Estimate (β)	Standard error	z-value	p-value	Estimate (β)	Standard error	z-value	p-value
Intercept	-0.42	1.28	-0.33	0.744	-1.16	0.85	-1.36	0.17
Slaughterhouse	5.39	9.30	0.58	0.562	168.48	230.90	0.73	0.47
Market	7.20	2.47	2.92	0.003	-447.35	129.40	-3.46	<0.001
Dumpster	-0.34	0.28	-1.20	0.231	785.43	147.61	5.32	<0.001
Dumpsite	0.02	0.03	0.65	0.516	153.05	200.77	0.76	0.45
ρ	0.934			< 2.22e-16	0.92			< 2.22e-16
	LR test value: 17.953, p-value: 2.2646e-05				LR test value: 16.397, p-value: 5.136e-05			
	Nagelkerke pseudo-R-squared: 0.73167				Nagelkerke pseudo-R-squared: 0.85644			

Direct and indirect effects of spatial lag model were then generated for the interpretation (Darmofal 2015). Market density had positive statistically significant direct effects on the cholera incidence rates within a ward (β = 12.45). A 1 square km increase of the market density within the ward is associated with 12.45 increases of cholera incidence rates within the ward. Distance from the centre of the community to market has shown a significant negative direct effect on cholera incidence rates of the ward (β = -710.37). A 1 km increase of distance from the centre of the community to the market within the ward is associated with 710.37 decreases in cholera incidence rates within the ward. The significant positive direct effect of distance from the community to dumpster on cholera incidence rates was also observed (β = 1247.23. A 1 km increase of distance from the centre of the community to the dumpster within the ward is associated with 1247.23 increases of cholera incidence rates within the ward (Table 5).

Table 5: Direct and indirect effects of spatial lag model of community risk factors related to cholera incidence rates

Variable	Kernel density variables			Average distance variables		
	Direct effect (β)	Indirect effect (β)	Total effects (β)	Direct effect (β)	Indirect effect (β)	Total effects (β)
Slaughterhouse	9.31	72.81	82.12	267.53	1849.86	2117.40
Market	12.45*	97.32	109.77	-710.37*	-4911.87	-5622.24
Dumpster	-0.59	-4.57	-5.15	1247.23**	8623.99	9871.22
Dumpsite	0.04	0.30	0.33	243.03	1680.44	1923.47

**p-value < 0.01, p-value < 0.05.

The community risk factors which were associated with cholera incidence rates are not consistent throughout the municipality. Geographically Weighted Poisson model was generated to show the variations of significant factors within the municipality. The results in Table 6 show that, market density (M = 0.43) and distance from the centre of the community to the dumpster of the ward (210.49) had positive associations with cholera incidence rates. Distance from the centre of the community to the market (M = -54.175) had a negative association on the cholera incidence rates.

Table 6: Parameter estimates in geographically weighted Poisson analyses for significant variables

Variable	Mean	Standard deviation	Minimum	Maximum
Intercept	-0.053	0.80	-1.10	1.47
Market density	0.43	0.51	-0.22	1.83
Distance to market	-54.18	64.67	-133.48	132.38
Distance to Dumpster	210.49	155.46	147.63	426.53

A statistically significant positive association of market density and cholera incidence rates was only observed in Mazimbu ward (p < 0.05). This means that, when market density in a ward increases, it is associated with an increase of cholera incidence rate (Figure 4).

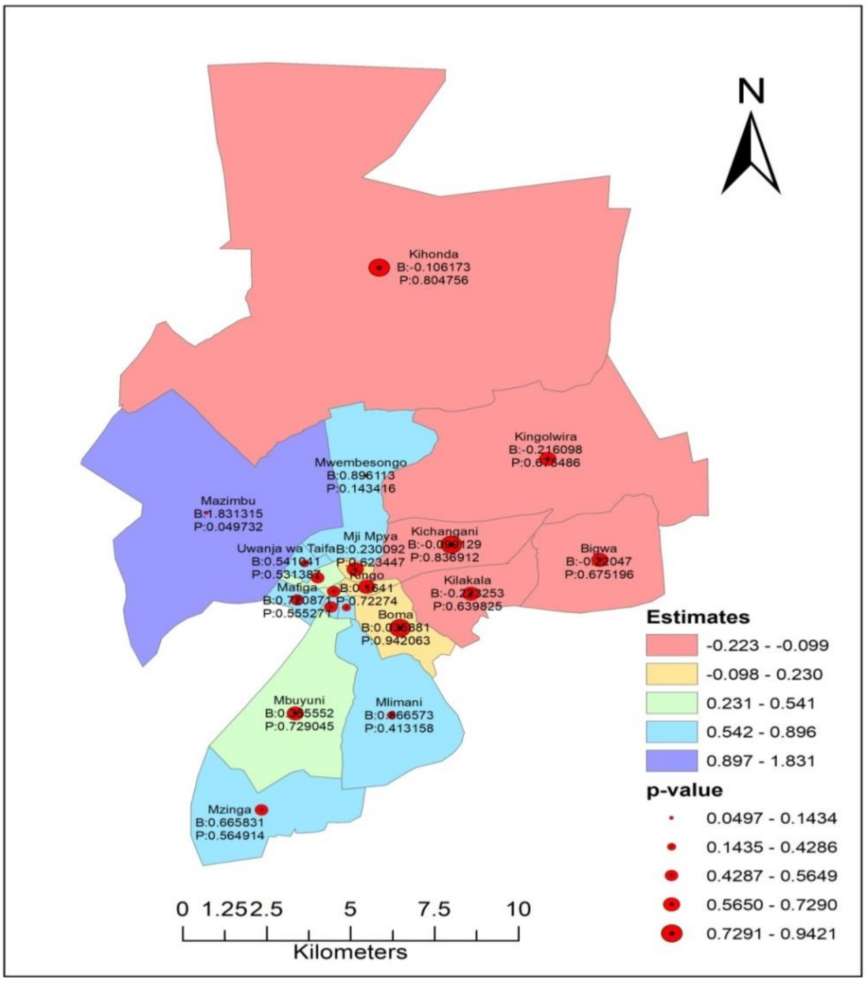

Figure 4: Spatial distributions of estimates and significant level of market density to cholera incidence rates.

Statistically significant negative associations of distance from the centre of the community to the market and cholera incidence rates was observed in Kihonda, Kingolwira and Kichangani ($p < 0.001$) and Bigwa wards ($p < 0.05$). This implies that, when the distance from the centre of the community to the market increased, the cholera incidence rate decreased (Figure 5).

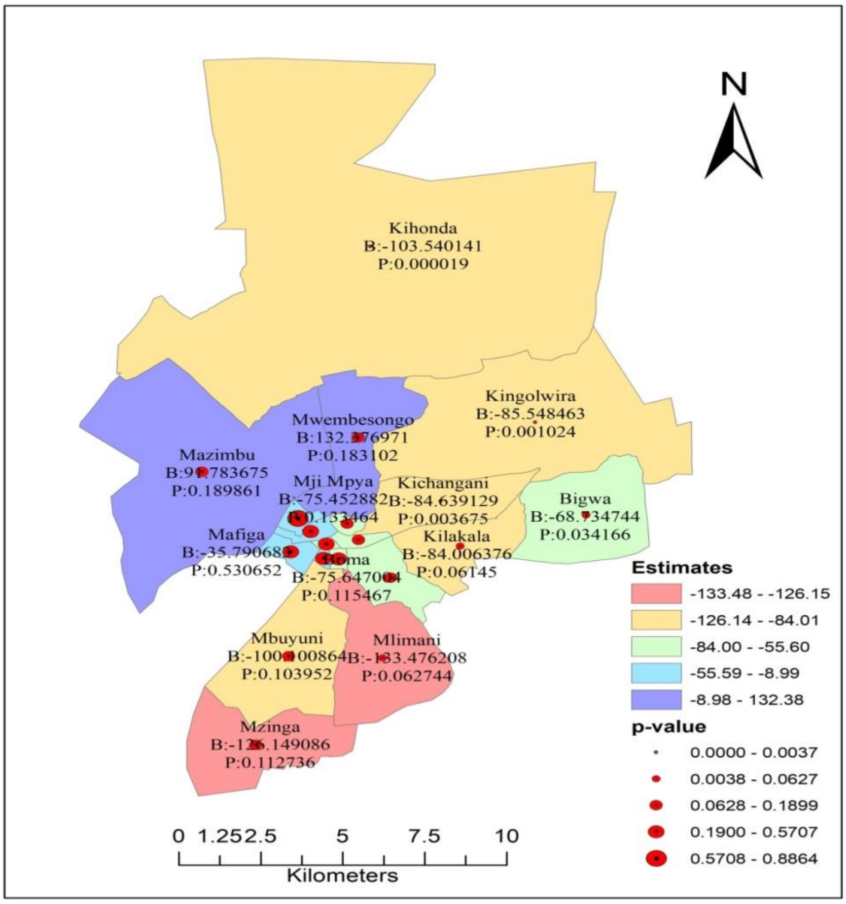

Figure 5: Spatial distributions of estimates and significant level of distance from the community to the market on cholera incidence rates.

Figure 6 shows statistically significant positive associations between distance from the community to the dumpster and cholera incidence rate in Kihonda, Kingolwira, Bigwa, Kichangani, Kilakala and Boma wards ($p < 0.001$) and some wards at the centre of the municipality which are Mji Mkuu and Kingo ($p < 0.05$).

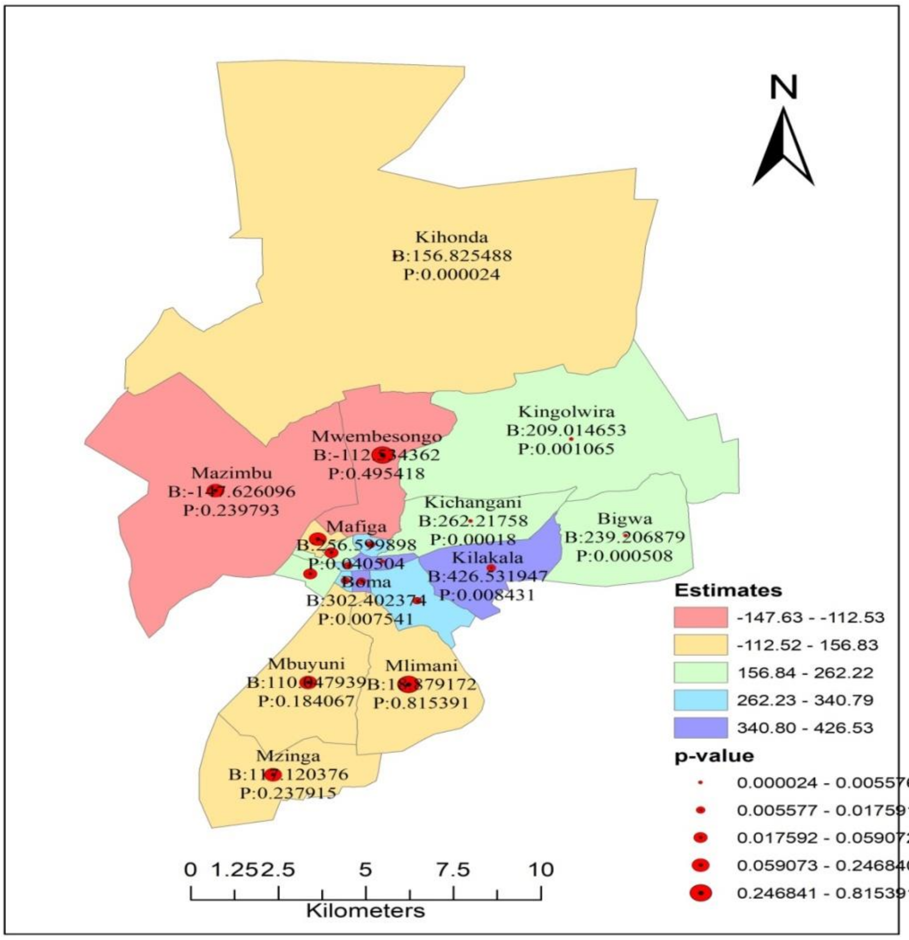

Figure 6: Spatial distributions of estimates and significant levels of distance from the community to the dumpster on cholera incidence rates.

Discussion

The community risk factors which are associated with cholera incidence rates are found not to be consistent throughout the wards within Morogoro municipality. This could be contributed by the fact that, the risk factors were not equally distributed in all wards. Geographically weighted Poisson analysis has shown the variations of significant explanatory variables of the spatial lag model. The findings of the current study have shown that market density is a significant predictor of cholera. This could be probably because there large amounts of wastes are produced in the markets Nyampundu et al. (2020) and waste management methods are among the hygiene practice components (Apate and Kamble 2019). Research conducted in Morogoro Municipality by Chengula et al. (2015) found unsatisfactory waste management services in the market areas. Less effective waste management and waste remaining for extended periods in the markets influence high bleeding sites of flies and other insects (Abdulrasoul and Bakari 2016, Nyampundu et al. 2020). Houseflies may help disseminate cholera by acting as mechanical vectors of the bacteria (Osei and Duker 2008). These flies may carry *V. cholerae* bacteria to the nearby households in different fomites which stay there for long hours and facilitate cholera transmission.

Unsafe methods of disposal of waste have effects on human health and survival. In urban areas, waste management is a serious public health problem since only 50% of the waste generated are effectively managed (URT 2017). It is likely for the disease to spread from its source to the nearby communities earlier than to those who live farther away (McKinley 2007, Schærström 2009). This study revealed that, communities near the markets were more likely to have higher cholera incidences than those far. This result is similar to findings of Olanrewaju and Adepoju (2017) who found proximity to the market and proximity to the waste dump site had relationships with cholera incidences. A study conducted in Ghana also recommended that cholera risk is high to the residents who live near to refuse dump and where there is high numbers of refuse dumps (Osei and Duker 2008). It is clear that cholera may spread through the intake of contaminated water or food whether by contact with an infected person or by direct/indirect contact with infected human faeces. The solid wastes accommodate different wastes; therefore, is very likely to meet human faeces in solid wastes through diapers or children faeces directly dumped to the temporary dumpsite (market) by the vendors. *Vibrio cholerae* can survive well in faeces if kept moist (Sack et al. 2004). Also, living near unmanaged solid waste is very likely to be exposed to diseases (Yoada et al. 2014).

Living in an area with no solid waste collection container is dangerous to the residents. The current study found that, communities that live far from the dumpster were more likely to have higher cholera incidence rates. There is a relationship between the distance of waste container located at the community and solid waste management (Odonkor et al. 2020). Waste collection method; like the provision of dumpster appear most dominant in the African countries for the purpose of dropping solid wastes from the nearby community (Anaezi et al. 2019, Lloyd 2019). It is also evident that lack of good and enough infrastructures leads to improper solid waste disposal in most urban centres (Srivastava et al. 2015, Chengula et al. 2015). The suitable disposal method of domestic wastes is by putting them into a closed plastic bin or nearby dumpster. However, most residents store solid waste in open bags and put it outside their homes which was found by Chengula et al. (2015) in Morogoro municipality and Yoada et al. (2014) in urban Accra due to lack of dumpsters. Even with improper disposal, delays of solid waste collection were also found in Morogoro municipality, and sometimes it took more than a month for the waste to be collected (Chengula et al. 2015). Unfortunately, improper disposal of solid wastes grants good breeding sites for flies which facilitate eruption and easy transmission of diarrhoeal diseases (Yoada et al. 2014, Odonkor et al. 2020), although those studies did not specifically aim cholera disease as one of the outcomes.

Conclusion

The current study shows that, risk factors associated with cholera vary across the wards. Market density, distance to the market and distance to the dumpster were the significant factors associated with cholera incidence rates in the wards. This means that market density, distance to the market and distance to the dumpster play significant roles in cholera transmission. The identified risk factors of cholera signify that education programmes and awareness on community waste management need to focus more on the urban areas where there is high production of waste. Also, timely collections of solid wastes and provision of dumpsters near the households to ensure that residents dispose their household wastes at known areas. However, those measures should be based on the variations of the risk factors found in the wards.

Acknowledgements

Authors provide gratitude to the Institute of Rural Development Planning (IRDP) for financial support and Morogoro Municipal office for provision of data and permission to use the data.

Conflict of Interest

The authors declare that there is no conflict of interest.

Research Permit

Research permit was obtained from the Vice Chancellor, University of Dar es Salaam.

References

Abdulrasoul AA and Bakari SS 2016 Challenges and problems of solid waste management in three main markets in Zanzibar. *Adv. Recy. Waste Manage*. 01(02): 1-9.

Ali M, Nelson AR, Lopez AL and Sack DA 2015 Updated global burden of cholera in endemic countries. *PLoS Neglect. Trop. Dis.* 9(6): e0003832.

Anaezi EL, Ozor PA and Nwanya SC 2019 Municipal solid waste collection and disposal using travelling salesman problem: A review. *Proceedings of the International Conference on Industrial Engineering and Operations Management* 2019(Mar): 1893-1903.

Anselin L 1988 *Spatial Econometrics: Methods and Models*. Springer Science & Business Media: Dordrecht, The Netherlands.

Apate DS and Kamble MN 2019 An analytical study of hygiene practices and its impact of health status of the household. *Int. J. Trend Sci. Res. Dev.* Special Issue-FIIIIPM2019): 152-159.

Baig MA 2017 *Using geographical information systems in epidemiology: mapping & analyzing occurrence of diarrhea in urban–residential area of Islamabad, Pakistan*. MSc Thesis, Lund. University, Sweden.

Chengula A, Lucas BK and Mzula A 2015 Assessing the awareness , knowledge , attitude and practice of the community towards solid waste disposal and identifying the threats and extent of bacteria in the solid waste disposal sites in Morogoro Municipality in Tanzania. *J. Biol. Agric. Healthcare* 5(3): 54-65.

Clemens JD, Nair GB, Ahmed T, Qadri F and Holmgren J 2017 Cholera. *Lancet* 390(10101): 1539-1549.

Cowman G, Otipo S, Njeru I, Achia T, Thirumurthy H, Bartram J and Kioko J 2017 Factors associated with cholera in Kenya, 2008-2013. *Pan Afr. Med. J.* 28(1): 156-156.

Daoud JI 2017 Multicollinearity and regression analysis. *Journal of Physics: Conference Series* 949(1): 012009.

Darmofal D 2015 *Spatial Analysis for the Social Sciences*. New York: Cambridge University Press.

Dureab F, Jahn A, Krisam J, Dureab A, Zain O, Al-Awlaqi S and Müller O 2019 Risk factors associated with the recent cholera outbreak in Yemen: a case-control study. *Epidemiol. Health* 41: e2019015.

Harris JB, LaRocque RC, Qadri F, Ryan ET and Calderwood SB 2012 Cholera. *Lancet* 379(9835): 2466-2476.

Hounmanou YM, Mølbak K, Kähler J, Mdegela RH, Olsen JE and Dalsgaard A 2019 Cholera hotspots and surveillance constraints contributing to recurrent epidemics in Tanzania. *BMC Res. Notes* 12(1): 1-6.

Lloyd JS 2019 Expanding safe waste management to public health systems. *The Lancet* 393(10168): 225.

McKinley TJ 2007 *Spatial survival analysis of infectious animal diseases*. PhD Thesis, University of Exeter.

Mengel MA, Delrieu I, Heyerdahl L and Gessner BD 2014 Cholera outbreaks in Africa. In: Nair G, Takeda Y (Eds) Cholera Outbreaks. *Current Topics in Microbiology and Immunology* 379:117-144. Springer, Berlin, Heidelberg.

Mollel EL 2016 *Economic analysis of solid waste management options in Morogoro Municipality, Tanzania*. MSc dissertation, Sokoine University of Agriculture.

NBS 2012 Population and Housing Census of Tanzania: Shapefiles- Level Three. Retrieved from https://www.nbs.go.tz/index.php/en/census-surveys/gis/386-2012-phc-shapefiles-level-three

Nyampundu K, Mwegoha WJ and Millanzi WC 2020 Sustainable solid waste management Measures in Tanzania: An exploratory descriptive case study among vendors at Majengo market in Dodoma City. *BMC*

Public Health 20(1): 1-16.
Odonkor ST, Frimpong K and Kurantin N 2020 An assessment of house-hold solid waste management in a large Ghanaian district. *Heliyon* 6(1): e03040.
Olanrewaju OE and Adepoju KA 2017 Geospatial assessment of cholera in a rapidly urbanizing environment. *J. Environ. Public Health* 2017.
Osei FB and Duker AA 2008 Spatial dependency of *V. cholera* prevalence on open space refuse dumps in Kumasi, Ghana: A spatial statistical modelling. *Int. J. Health Geographics* 7: 1-17.
Sack DA, Sack RB, Nair GB and Siddique AK 2004 Cholera. *Lancet* 363(9404): 223–233.
Schærström A 2009 Disease Diffusion. *International Encyclopedia of Human Geography* 222-233.
Silverman BW 1986 *Density estimation for statistics and data analysis.* London; New York: Chapman and Hall.
Srivastava V, Ismail SA, Singh P and Singh RP 2015 Urban solid waste management in the developing world with emphasis on India: challenges and opportunities. *Rev. Environ. Sci. Biotechnol.* 14(2): 317-337.
Steven S and Mkonda MY 2015 Climate variability impacts, diseases and human health: the case of Morogoro Municipalities, Tanzania. *Archives of Business Research* 3(5): 78-90.
Trærup SL, Ortiz RA and Markandya A 2011 The costs of climate change: A study of cholera in Tanzania. *Int. J. Environ. Res. Public Health* 8(12): 4386-4405.
Unhabitat 2009 *Tanzania: Morogoro Profile. Regional and Technical Cooperation Division, Nairobi*.
UNICEF 2018 *Tanzania Humanitarian Situation Report*. http://www.unicef.org/ethiopia/UNICEF_Ethiopia_Humanitarian_SitRep_January_2016.pdf
URT 2014 Water Sector Development Programme Phase II. In *Dar es Salaam: Ministry of Water, United Republic of Tanzania.*
URT 2015 *Migration and Urbanization Monograph*. *IV*, 1–68.
URT and ICF 2016 *Tanzania Demographic and Health Survey and Malaria Indicator Survey* (TDHS-MIS) 2015-16. Dar es Salaam, Tanzania, and Rockville, Maryland, USA: MoHCDGEC, MoH, NBS, OCGS, and ICF.
URT 2017 Policy–National Health Policy. *Ministry of Health, Community Development, Gender, Elderly, and Children, Tanzania*, *October*, 1-55.
Wahed T, Kaukab SS, Saha NC, Khan IA, Khanam F, Chowdhury F, Saha A, Khan AI, Siddik AU, Cravioto A, Qadri F and Uddin J 2013 Knowledge of, attitudes toward, and preventive practices relating to cholera and oral cholera vaccine among urban high-risk groups: Findings of a cross-sectional study in Dhaka, Bangladesh. *BMC Public Health* 13:242.
Yoada RM, Chirawurah D and Adongo PB 2014 Domestic waste disposal practice and perceptions of private sector waste management in urban Accra. *BMC Public Health* 14(1): 1-10.
You A, Ali M, Kanungo S, Sah B, Manna B, Puri M, Nair GB, Bhattacharya SK, Convertino M, Deen JL, Lopez AL, Wierzba TF, Clemens J and Sur D 2013 Risk map of cholera infection for vaccine deployment: The Eastern Kolkata Case. *PLoS One* 8(8): e71173.