Spatial Heterogeneity and Determinants of Diarrhoea among Under-Five in Pakistan

Keywords: Diarrhoea, Socio-demographic factors, Environmental condition, Children under five, MGWR, Pakistan

Munazza Fatima¹ and Ömer Ünsal²

¹Department of Geography, The Islamia University of Bahawalpur, 63100 Punjab, Pakistan <u>munazza.fatima@iub.edu.pk</u>

²Institute of Social Sciences, Istanbul University, 34116 Istanbul, Turkey, <u>oomer.unsal@gmail.com</u>

Background

- Diarrhoea is the second leading cause of death under-five
- 1.7 billion cases & <u>525,000 deaths</u> ANNUALY (Bristol=483,000 total pop*)
- Deaths=53,000 in Pakistan

* https://www.bristol.gov.uk/council/statistics-census-information/population-of-bristol#:~:text=Mid-2023%20population%20estimates%20were,estimated%20to%20be%20483%2C000%20people.

Purpose

 This study aims to spotlight geographical differences in the determinants of under-five diarrhoea cases across districts in Pakistan using a cross-sectional approach.

Study Area: Pakistan

- It ranks as the world's 5th most populous country, with a current population of 231 million people and a population growth rate of 2 percent per year.
- The country has 32.2 million households, with 12.2 million located in urban areas and 20 million in rural areas.
- Pakistan's literacy rate is 62.3%, leaving roughly **60 million** people in the country illiterate.
- Approximately **25 million** people still engage in open defecation
- An estimated **70%** of households still drink bacterially contaminated water
- Almost **20%** of households lack basic hand hygiene facilities

Data

Variables		Category	Variable Name	Measure and Unit	Data Source
Outcome		Disease	Diarrhoea among <5 year	Percentage of children aged 0-59 months for whom the mother/caretaker	MICS*
				reported episodes of diarrhoea during the last two weeks	
Explanatory	V1	Demography	Population Density	Persons/square kilometer	Census 2017
	V2		Rural Population	Percentage of Rural Population	Census 2017
	V3	Health	Episodes of Fever	Percentage of children aged 0-59 months for whom the mother/caretaker	MICS*
				reported fever in the last two weeks	
	V4		ARI	Percentage of children aged 0-59 months for whom the mother/caretaker	
				reported symptoms of ARI and/or fever in the last two weeks	
	V5		Premature birth	Percent distribution of pre-mature births (birth before 37 weeks of pregnancy)	
				in last 2 year	

Data

Variables		Category	Variable Name	Measure and Unit	Data Source
Explanatory	V6	Sanitation	Open stool	Percentage of children whose last stools were Left in the open space	MICS*
	V7	Nutrition	Children's vitamin A	Percent distribution of children aged 6-59 months by receipt of a high dose	
			supplementation	vitamin A supplement in the last 6 months	
	V8		Salt Iodisation	Percentage of households (By consumption of iodized salt)	
	V9	Mother	Breast Feed	Children aged 0-23 months Percent appropriately breastfed	
		attributes			
	V10	Meteor	Mean Temperature of	Average annual temperature	World Clima
		Meteorological	Warmest Quarter		(www.worldclim
					.org)









• The first Stage : Data processing for analysis

38 different variables that may affect under-five diarrhoea cases were brought together at the district level (the geographical sub-unit of the province).



• In the second stage,

Generalized Linear Regression (GLR), which is a global regression analysis method, was used. In this tool, the model type was determined as Continuous (Gaussian). This method reveals the direction (positive, negative) of the relationship between the dependent variable and independent variables.

•
$$y_i = \beta_0 + b_{1x1} + b_{2x2} + b_{nxn} + e$$



 In the third stage, the Spatial Autocorrelation (Global Moran's I) tool was used to test the presence of statistically significant clustering in the linear regression residuals.



- In the fourth stage, the Multiscale Geographic Weighted Regression (MGWR) method was used to overcome the clustering problem that occurs in the GLR model residuals, to improve the predictions, and to understand regional variation and relationship direction in the study area.
- GWR, Similarity GWR and MGWR results were compared.
- In the last step, bivariate maps were produced to better evaluate the relationship between MGWR beta values and the explanatory variable.

Limitations

- First, it was not possible to do a comparison study because there was **no prior study** available on the spatial trends of diarrhoea in Pakistan.
- Second, the analysis was conducted at the district level due to the **absence of tehsil-level** (the smallest spatial unit of the municipality) data, limiting our ability to explore more localized variations.
- Third, the data used in this research was collected during **different years** throughout the provinces via MICS and lacks temporal disease data.

Results of the GLR Model

GLR Results	Value	GLR Results	Value
Number of observations	160	Akaike Information Criterion (AICc)	1040.825352
Multiple R-Squared	0.740702	Adjusted R-Squared	0.723299
Joint Statistic	42.562819	Prob(>F), (10,149) degrees of freedom	0.000000*
Joint Wald Statistic	489.433	Prob(>chi-squared), (10) degrees of freedom	0.000000*
Koenker (EP) Statistic	19.615688	Prob(>chi-squared), (10) degrees of freedom	0.033104*
Jarque-Bera Statistic	70.659167	Prob(>chi-squared), (2) degrees of freedom	0.000000*

These values indicate that the GLR regression residuals have a normal distribution.

Results of the GLR Model

Variable	Coefficient	Std. Error	t-Statistic	Probability	Robust SE	Robust t	Robust Prb	VIFC
Intercept	-5.767835	3.192188	-1.80690	0.072799*	3.223797	-1.78914	0.075628*	
Children-Vitamin A	-0.063576	0.021941	-2.89764	0.004330*	0.020807	-3.05558	0.002668*	1.902552
Breastfeed	0.095243	0.035651	2.67154	0.008386*	0.039067	2.437958	0.015938*	1.809951
Population Density	-0.000001	0.000000	-2.473,23	0.014505*	0.000000	-2.308,87	0.015938*	1.493408
Open stool	-0.13675	0.038488	-3.552.91	0.000520*	0.035644	-3.836.39	0.000191*	1.226731
Mean Temperature of Warmest Quarter	0.208718	0.095936	2.175602	0.031153*	0.10355	2.015623	0.045633*	1.391451
Pre- Mature Birth	0.072647	0.027603	2631838	0.009381*	0.030395	2.390,141	0.018082*	1.100924
Salt -iodization	-0.054505	0.021596	-2523778	0.012651*	0.021751	-2.505,87	0.013282*	1.895553
ARI	0.478241	0.107192	4.461,521	0.000018*	0.133362	3.586030	0.008464*	2.276789
Fever Episodes	0.549092	0.066505	8.256367	0.000000*	0.074228	7.397377	0.000000*	3.244894
Rural Population	0.000005	0.000001	4.186211	0.000053*	0.000001	4.380322	0.000025*	1.352425

Clustering of GLR Model Residuals

- When Moran's Index is applied to ensure that the residuals of the GLR model are spatially random, we find that the residuals for diarrhoea prevalence of under five **are spatially clustered**.
- Diarrhoea prevalence of under-five was found to be clustered at the district level with a **Moran's Index value of 0.09.** This finding is statistically significant at a 99% level with **z-score of 3.02 and a p-value of 0.0.**

Results of the GWR-SGWR-MGWR Models

Model Diagnostics			
Statistic	GWR	MGWR	SGWR
R- Squared	0.8320	0.8653	0.88
Adjusted R- Squared	0.7947	0.8628	0.842
AIC _c	228.1595	213.648	225.711
Sigma Squared	0.2050	0.1670	0.397
Sigma Squared MLE	0.1680	0.1347	
Effective Degrees of Freedom	131.1603	129.0926	120.572

The SGWR (a python library now, it doesn't have a GUI or extension for ArcGIS or QGIS) didn't calculate p-values of the results by default. Also it can work with only point geometry, now.

We will share with you the findings of **the MGWR model** instead of GWR or SGWR due to the lowest AICc and highest Adjusted R2 values.

Results of the MGWR Model

Explanatory Variables	Neighbors (% of Features) ^a	Significance (% of features) ^b
Intercept (Scaled)	160(100.00)	14(8.75)
Children-Vitamin A (Scaled)	61(38.12)	11(6.88)
Breastfeed (Scaled)	141 (88.12)	100(62.50)
Population Density (Scaled)	62(38.7S)	78 (48.7S)
Open stool (Scaled)	160(100.00)	93 (58.13)
Mean Temperature of Warmest Quarter (Scaled)	118 (73.375)	57 (35.52)
Pre- Mature Birth (Scaled)	160(100.00)	9(5.62)
Salt-iodization (Scaled)	94(54.75)	86(5.75)
ARI (Scaled)	63 (39.38)	110(68.75)
Eever Episodes (Scaled)	72 (45.00)	160(100.00)
Rural Population (Scaled)	129 (9.62)	98 (61.25)

a: This number in the parenthesis ranges from 0 to 100% and can be intercepted on a local, regional, or global scale based on the geographical context from low to high.

b: In the parentheses, the percentage of features that have significant coefficients of explanatory variables.

Results of the MGWR Model-Goodness of Fit



- The results showed that the local R2 goodness of fit effect was good in most of the districts of Pakistan.
- 87.5% of Pakistan districts have a goodness of fit exceeding 0.80.

Results of the MGWR Model



Spatial distribution of coefficients MGWR model results for episodes of Diarrhoea among children in Pakistan. This map is also showing **significant and non-significant areas**.

V10 Mean Temp. of Warmest Q.

Bivariate Maps of the MGWR Model Coefficients and Variable Values



Spatial distribution bivariate map of local estimates (Beta) and covariates original values of MGWR model results for Diarrhoea episodes among children in Pakistan

Discussion

Significant Diarrhoea-Associated Variables Revealed by GLR

- The relationships between diarrhoea rates in children under five years of age and vitamin A supplement, salt-iodization, temperature, temperature, premature birth, ARI, episode of fever, and rural population are <u>consistent</u> with the literature.
- The relationships between diarrhoea rates in children under five years of age and open stool, population density, breastfeed are <u>unconsistent</u> with the literature.

Discussion

GLR Model

- Open defecation, vitamin A supplementation, and salt iodization reduce diarrhoea prevalence.
- These factors have been shown to reduce diarrhoea prevalence, aligning with global research on the protective effects of nutrition and improved sanitation practices.
- Breastfeeding, higher temperatures, premature births, ARI, fever, and rural population increase diarrhoea risk.
- Warmer temperatures increase pathogen growth; poor breastfeeding practices and rural living conditions expose children to higher risks of infection.

Discussion

MGWR Model

- Different districts show varying relationships with diarrhoea determinants.
- Central districts have a strong link between premature births and diarrhoea.
- Northeastern districts have a negative association with salt iodization, indicating regional disparities in nutritional content.
- Breastfeeding is a significant factor in diarrhoea prevalence in over half the country.

Policy Recommendations

- Enhance Nutritional Programs
- Expand Vaccination Coverage
- Improve Sanitation and Water Access

Future Research Directions

- Conduct **longitudinal studies** to evaluate the long-term impact of targeted interventions in different regions of Pakistan
- Explore additional socio-environmental factors using non-linear models to capture complex relationships in Pakistan's diverse geographical and socio-economic contexts

Conclusion

- Diarrhoea in children under five varies significantly across Pakistan due to local factors like nutrition, environment, and socioeconomics
- Target interventions in high-risk areas should include immunization, better sanitation, clean water, and nutrition

 Conduct longitudinal studies and use advanced spatial models to refine strategies and policies

Thanks

Spatial Heterogeneity and Determinants of Diarrhoea among Under-Five in Pakistan

Munazza Fatima¹ and Ömer Ünsal²

¹Department of Geography, The Islamia University of Bahawalpur, 63100 Punjab, Pakistan <u>munazza.fatima@iub.edu.pk</u> ²Institute of Social Sciences, Istanbul University, 34116 Istanbul, Turkey, <u>oomer.unsal@gmail.com</u>