

SCIENTIFIC REPORTS



OPEN

Spatio-temporal analysis of small-area intestinal parasites infections in Ghana

F. B. Osei^{1,2} & A. Stein²

Intestinal parasites infection is a major public health burden in low and middle-income countries. In Ghana, it is amongst the top five morbidities. In order to optimize scarce resources, reliable information on its geographical distribution is needed to guide periodic mass drug administration to populations of high risk. We analyzed district level morbidities of intestinal parasites between 2010 and 2014 using exploratory spatial analysis and geostatistics. We found a significantly positive Moran's Index of spatial autocorrelation for each year, suggesting that adjoining districts have similar risk levels. Using local Moran's Index, we found high-high clusters extending towards the Guinea and Sudan Savannah ecological zones, whereas low-low clusters extended within the semi-deciduous forest and transitional ecological zones. Variograms indicated that local and regional scale risk factors modulate the variation of intestinal parasites. Poisson kriging maps showed smoothed spatially varied distribution of intestinal parasites risk. These emphasize the need for a follow-up investigation into the exact determining factors modulating the observed patterns. The findings also underscored the potential of exploratory spatial analysis and geostatistics as tools for visualizing the spatial distribution of small area intestinal worms infections.

Intestinal parasites, helminth and protozoa, are among the most common infections of humans in developing countries with considerable morbidity and substantial burden on public health. The common intestinal parasites, *Ascaris lumbricoides* (roundworms), *Trichuris trichiura* (whipworms), and *Necator americanus* or *Ancylostoma duodenale* (hookworms) are estimated to cause one fourth of the known human infectious diseases¹. Although there are concerns about the actual numbers of infections², intestinal parasites are estimated to infect more than 1 billion people^{3,4}. In children, infections can retard child growth, cause anemia, and create cognitive and physical challenges⁵⁻⁷. Prevalence is high amongst resource-poor countries, especially in Asia, Sub-Saharan Africa, and Latin America where there is limited access to water supply and poor sanitation. Chemotherapy based on Mass Drug Administration (MDA) is the contemporary control strategy embraced by the World Health Organization (WHO) and governments to reduce infections. The recommendation is to ensure periodic administration of albendazole and mebendazole to at-risk populations^{3,8}. To ensure that MDA control programs are focused appropriately on efficient utilization of scarce resources, targeting of populations requiring intervention is essential. This requires an understanding of the nature of spatial patterns, and precise estimates of the local risks for comparison.

Ghana is among the Sub-Saharan countries with high prevalence of intestinal parasites infections^{3,4}. The disease has constantly been listed amongst the top five outpatient morbidities. Prevalence has been reported to be between 2% and 78% for various parasites and specific population groups⁹⁻¹⁴. Using random sampling, Ayeh-Kumi *et al.*¹⁰ estimated the prevalence of intestinal parasites among food vendors in Accra Metropolitan area to be 21.6%, with helminthic predominating with 15.2%. The prevalence of hookworm infection among school-aged children was estimated to be 39.1% in the Kintampo North Municipality, with significant risk factors being age, malaria parasitemia, lack of health care, school area, levels of antibodies against hookworm, and low consumption of animal foods. Asymptomatic carriage among psychiatric patients was estimated to be 13.5%¹³ among some Ghanaian orphanages. These studies focused on either the biological or anthropogenic characteristics of the individuals affected. Besides, prevalences have either been estimated for single geographic units

¹Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, Ghana.

²Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands. Correspondence and requests for materials should be addressed to F.B.O. (email: f.b.osei@utwente.nl)

or among specific population categories, hence are unable to evaluate the spatial patterns of infection. Intestinal parasites thrive under climatic and environmental conditions such as warm temperatures, high precipitation and adequate soil moisture^{15–17}. Infections have often been associated with sociodemographic conditions such as poverty, poor sanitation, and poor drinking water^{18–20}. In Ghana, prevalence has been associated with sociodemographic conditions¹⁴. Since these underlying risk factors are spatially dependent, morbidity rates will be expected to exhibit spatially dependent patterns.

Spatial analysis and geostatistics can provide opportunities to study the type and nature of spatial patterns, and where these patterns occur. They have widely been used to study the spatial patterns and estimate the spatial risk of intestinal parasites infections^{17,21–25}. Spatial analysis methods such as the global Moran's Index²⁶ (Moran's I hereafter) and its local counterpart, Anselin's Local Indicator for Spatial Association (LISA)²⁷, could illuminate potential causal factors of diseases^{28,29}. Geostatistical analysis of health outcomes has also recently received increasing attention as a filtering tool^{30–32}. For instance, Poisson kriging allows filtering of those noise by integrating population heterogeneities to account for non-constant variance³⁰.

In this paper, we utilize spatial analysis tools and geostatistics to study the spatial patterns and provide spatially explicit maps of risk estimates useful for guiding control programs. Our specific objectives are to (1) quantify the type of spatial association, and detect and map clusters, (2) quantify the nature of spatial structure and map risk estimates of intestinal parasites using district level morbidities in Ghana. As neighborhood health planning in Ghana is largely based upon small-areas (administrative districts), studying spatial patterns of infections at the district level will present valuable and easy to implement information.

Methods

Study area and data. *Study Area and Data.* Ghana is a tropical region centrally located on the West Coast of Africa with a total land area of 239,000 km² (Fig. 1: created with ArcGIS software). The average annual temperature is approximately 26 °C (79 °F). There are two distinct rainy seasons, April–June and September–November, but March–September for the northern belt. Annual rainfall ranges from 1,015 mm in the north to 2,030 mm in the southwest (PHC, 2010). The country consists of ten administrative regions which are subdivided into 216 districts. Ghana is subdivided into six agro-ecological zones: Sudan Savannah, Guinea Savannah, Coastal Savannah, Forest/Savannah transitional zone, Deciduous Forest zone and the Rain Forest zone (Fig. 2).

In this study, we used aggregated clinically or laboratory diagnosed cases of intestinal worms parasites. Due to the protection of patient privacy and perhaps deficiencies in address geocoding systems, publicly available data on precise locations of disease cases are uncommon. Consequently, the spatial scale of our study was limited to the 170 administrative districts of which data were available. We obtained district level yearly aggregated cases of intestinal parasites infections from 2010 to 2014 from the Centre for Health Information and Management (CHIM) of the Ghana Health Service (GHS). CHIM is responsible for compiling and ensuring uniformity in reporting and managing all morbidities reported to health facilities (clinics and polyclinics, hospitals). In summary, health facilities capture and aggregate data and submit to sub-districts. Sub-districts aggregate facility summary reports and submit those to districts. Districts then receive both facility and sub-district summary reports for validation. At each stage of the data recording hierarchy, data can be entered directly into the District Health Information Management System (DHIMS). We also obtained population estimates for 2010 to 2014 from the Ghana Statistical Service (GSS).

Spatial autocorrelation. We used both global and local Moran's I of spatial autocorrelation to estimate the strength of spatial correlation. The global Moran's I ²⁶ estimates the general strength of spatial autocorrelation among districts while its local equivalent, LISA²⁷, estimates the spatial autocorrelation between districts and their neighboring districts. Thus, the local Moran's I identifies districts with high and low risks as well spatial outliers. For the observed counts y_i and populations n_i for the set of districts $i = 1, \dots, m$, we assumed that the counts are realizations from the Poisson distribution $Y_i | r_i \sim \text{Poisson}(n_i \cdot r_i)$. The variable of interest is the risk r_i ; its maximum likelihood estimate equals $r_i = y_i/n_i$, and its variance $v_i \propto n_i^{-1}$. Since the variance is inversely proportional to the population sizes, the required assumption of constant variance is violated and could lead to misleading results of large variances for regions with small populations. We used the empirical Bayesian standardization to account for the unequal variances arising from unequal populations³³. To do so, we constructed a standardized variate $r_i^{EB} = (r_i - \bar{r})/\sqrt{v_i}$ based upon the unconditional marginal expectation \bar{r} and district specific variance $v_i = s^2 + \bar{r}/n_i$. The method of moments estimates^{34,35} for the mean and variance equals $\bar{r} = \Sigma y_i/\Sigma n_i$ and $s^2 = \sigma^2 - \bar{r}/(\Sigma n_i/m)$, respectively, where $\sigma^2 = \Sigma n_i(r_i - \bar{r})^2/\Sigma n_i$.

We computed the global Moran's I , $I = \Sigma_i \Sigma_j w_{ij} r_i^{EB} r_j^{EB}$, and local Moran's I , $I_i = r_i^{EB} \Sigma_j w_{ij} r_j^{EB}$ using the standardized variable z_i . We defined the binary connectivity weight matrix w_{ij} as $w_{ij} = 1$ if $(i) \cap (j) \neq \text{null}$, and 0 otherwise, where (i) and (j) are the set of boundary points of district i and j , respectively. The m by m weight matrix w_{ij} was row-standardized, and satisfied the following conditions (1) symmetry, i.e., $w_{ij} = w_{ji}$, (2) zero diagonal elements, i.e. $w_{ii} = 0$, and (3) normalization, i.e., $\Sigma_j w_{ij} = 1$. For the local index, I_i , the summation over j implies that only the set of neighbors J_i of i , $j \in J_i$, was included.

To test the null hypothesis of no spatial autocorrelation, we generated 999 independent permutations of the vector (z_1, \dots, z_m) and computed I and I_i for each permuted vector to generate the empirical distribution. The p -value was estimated as the proportion of the number of times the index from the permuted data exceeds the Index from the actual data. We used the *spdep*³⁶ package of the R statistical software³⁷ for estimating both global and local Moran's I .

Spatial structure and smoothing. We used geostatistical smoothing to filter out the noise caused by heterogeneous population distribution. Unlike deterministic smoothers³⁸, geostatistical smoothing can account for

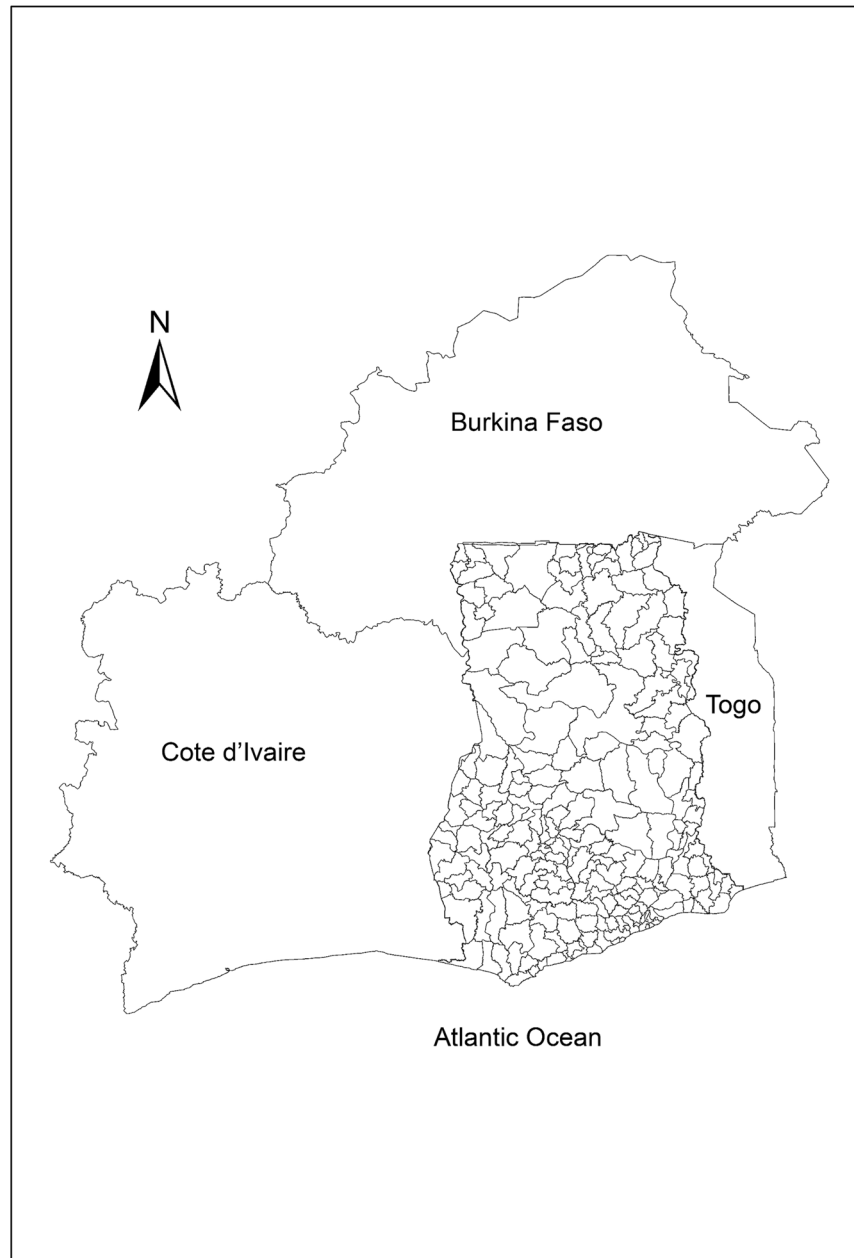


Figure 1. District map of Ghana showing its neighboring countries; Cote d'Ivoire (left), Burkina Faso (top) and Togo (right). This map was created using ArcGIS software (version 10.1, ESRI Inc. Redlands, CA, USA. <https://www.esri.com/>).

the range of spatial correlation and estimate the associated uncertainties. We assumed that the risks r_i are realizations from a second-order stationary random field. Poisson kriging was used to estimate the risk over a given district i_0 as linear combinations of the risk observed for that district r_i , and its neighboring districts $\hat{r}_{i_0} = \hat{\mu} + C_{i_0} C_{ij}^{-1} (r_i - \hat{\mu} \mathbf{1}_m)$. The assumption of stationarity implies the spatial mean of the prediction locations $\hat{\mu}_{i_0}$ is the same as the spatial mean of the random variable r_i , $E(r_i) = \hat{\mu}_{i_0} = \hat{\mu}_i$. Under minimum variance, the best linear unbiased estimate of the spatial mean equals $(\mathbf{1}'_m C_{ij}^{-1} \mathbf{1}_m)^{-1} \mathbf{1}'_m C_{ij} r_i$. We refer to the vector $(r_i - \hat{\mu} \mathbf{1}_m)$ as the predictor variables, C_{i_0} as a covariance vector between the prediction location and the predictor variable, and C_{ij} as the covariance matrix of the predictor variables. Essentially, $C_{i_0} C_{ij}^{-1} = \lambda_{i_0}$ yields the so-called kriging weights. We used the variogram model $\hat{\gamma}(h)$ as a structural tool to estimate the covariance function based on the relation $C_{ij} = C_{ii} - \hat{\gamma}(h)$, where C_{ii} is the variance of the risk or covariance at lag 0.

We used the empirical variogram estimator $\gamma(h) = \{2 \Sigma w_{ij}\}^{-1} \Sigma \{w_{ij} (r_i - r_j)^2 - r^*\}$, where $h = |i - j|$ and $N(h)$ is the number of observation pairs separated by the distance h between the centroids of districts i and j . Here, the variogram $\gamma(h)$ depends only on the distance between the centroids of districts, and this refers to the assumption of uniform population density within each district. This is an adjusted experimental variogram estimator

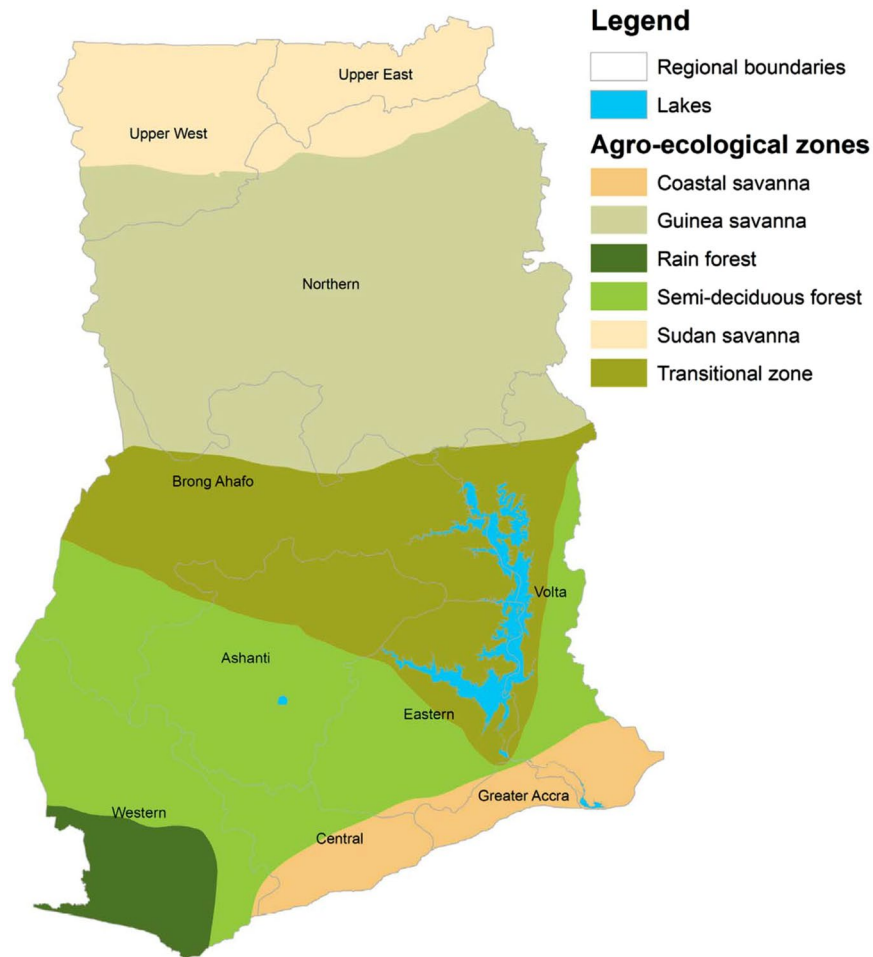


Figure 2. Map showing the six agro-ecological zones of Ghana. This map was created using ArcGIS software (version 10.1, ESRI Inc. Redlands, CA, USA. <https://www.esri.com/>).

proposed by^{39,40}, and generalized by Goovaerts³⁰ for disease mapping to account for heterogeneous populations. The rate differences ($r_i - r_j$) are weighted by their corresponding populations $w_{ij} = \frac{n_i \cdot n_j}{n_i + n_j}$ to homogenize their variances, where $r^* = \frac{\sum n_i r_i}{\sum n_i}$ is the population-weighted mean.

To explore both local $\gamma_{loc}(h)$ and regional $\gamma_{reg}(h)$ level spatial correlations, we fitted permissible nested variogram models $\hat{\gamma}(h) = \hat{\gamma}_{loc}(h) + \hat{\gamma}_{reg}(h)$ to $\gamma(h)$ ^{41–43}. This implies that the risk with mean μ was decomposed as the sum of local $r_{i,loc}$ and regional $R_{i,reg}$ orthogonal random functions, $r_i = r_{i,loc} + r_{i,reg} + \mu$, each with its particular contributory variogram $\hat{\gamma}_{loc}(h)$ and $\hat{\gamma}_{reg}(h)$, respectively. This is useful to unravel scale-dependent spatial autocorrelation patterns. The commonly used variogram models such as the exponential and spherical models have been described elsewhere^{43,44}. We fitted local and regional scale nested spherical models $\hat{\gamma}_{loc}(h)$ and $\hat{\gamma}_{reg}(h)$, respectively, using weighted least squares with weights $w_i(h) = \sqrt{N(h)}/\gamma(h)$. In this paper, the variogram modeling was conducted using the public-domain software `poisson_kriging.exe`³⁰.

Results

Spatial autocorrelation. Between 2010 and 2014, a total of 3,310,653 intestinal parasites infections were reported. The annual incidence rates ranged from 1.55% to 3.3%, with an average annual incidence rate of 2.53% (Table 1). The incidence rate increased from 1.55% in 2010 to 3.3% in 2014, with a slightly lower rate of 3.25% in 2014. We found significant positive spatial autocorrelations throughout 2010 to 2014 (Table 1), indicating that districts of similar risks were spatially clustered. Global autocorrelation was highest in 2010 ($I = 0.388$, $p = 0.01$) and lowest in 2012 ($I = 0.095$, $p = 0.01$). For easy interpretation, we presented the results of the local spatial autocorrelations as cluster maps based upon four categories: high-high, low-low, high-low, and low-high (Fig. 3). The high-high and the low-low associations indicate clustering of high risk (hot-spots) and low risks, respectively. Both the high-high and low-low associations indicate significant ($p \leq 0.05$) clustering of similar risks or positive spatial autocorrelation. The low-high category indicates that high risk districts surround a low risk district, whereas the high-low category indicates that low risk districts surround a high risk district. These are indications of spatial outliers. High-high clustering dominated within the middle belt while low-low clustering dominated within the northern parts. Although we undertook no formal causal relationships since this is an exploratory

Year	Cases	Rate (%)	Moran's <i>I</i>	<i>p</i> -value
2010	381307	1.55	0.388	0.01
2011	505614	2.00	0.367	0.01
2012	663576	2.57	0.095	0.01
2013	875921	3.30	0.248	0.01
2014	884232	3.25	0.165	0.01

Table 1. Yearly incidence rates and global Moran's *I*.

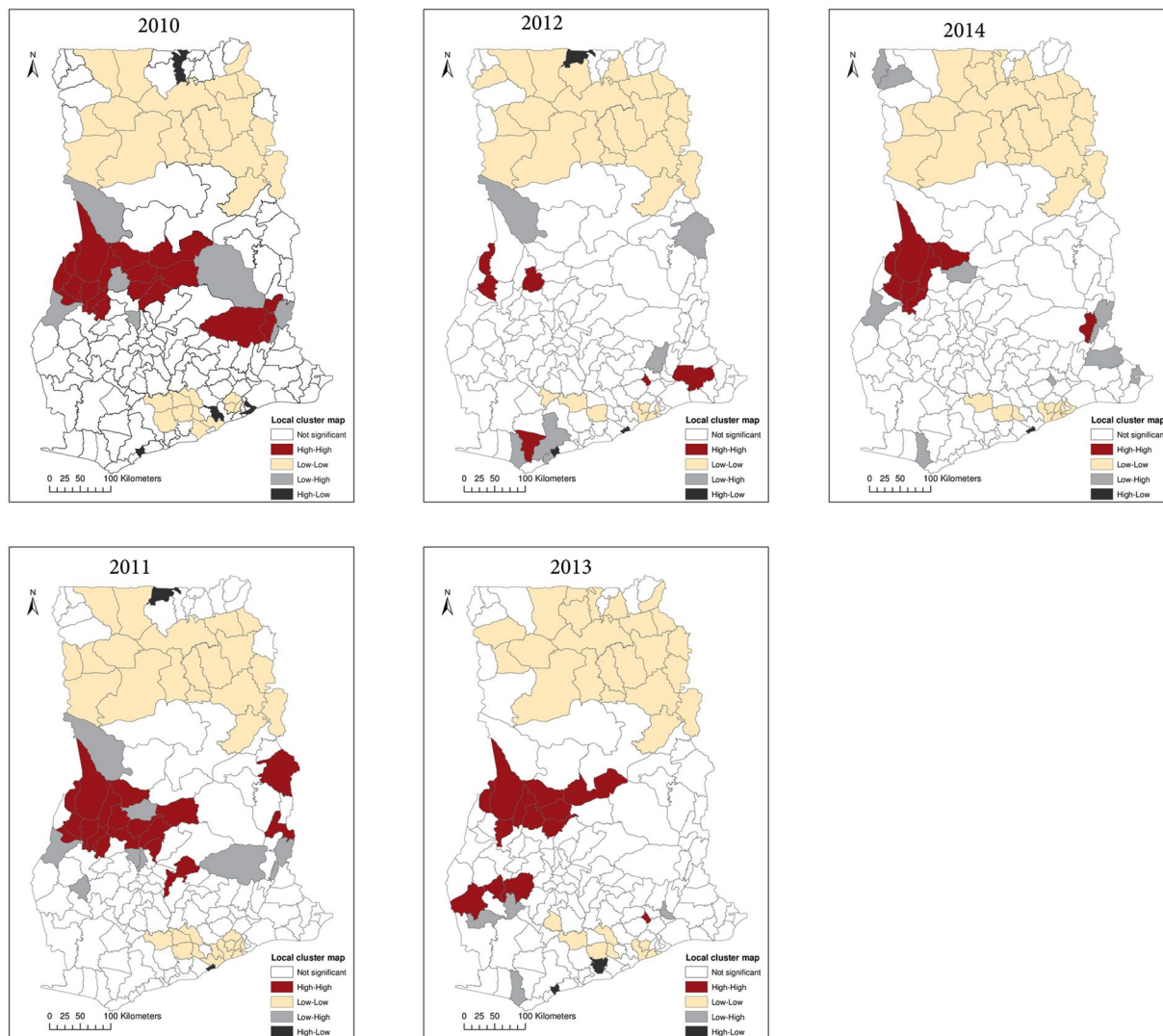


Figure 3. Local Moran's *I* cluster maps showing high-high, low-low, low-high, and high-low spatial associations. This map was created using ArcGIS software (version 10.1, ESRI Inc. Redlands, CA, USA. <https://www.esri.com/>).

study, a visual assessment of the local Moran's *I* maps together with the ecological zones map of Ghana (Fig. 2) was worthwhile. We found that the high-high clusters extended within the semi-deciduous forest and the transitional ecological zones. The low-low clusters on the other hand were concentrated within the northern parts and mostly intersected with the Guinea and Sudan Savannah ecological zones. Few outliers were detected throughout the study period.

Spatial structure and smoothing. We computed experimental variograms for each year using 10 km lag distances for 20 lags and fitted nested spherical models (Fig. 4). At 10 km lags, there were enough (≥ 30) pairs of districts to obtain stable estimates of the variogram. Table 2 shows the parameters of the models fitted to the

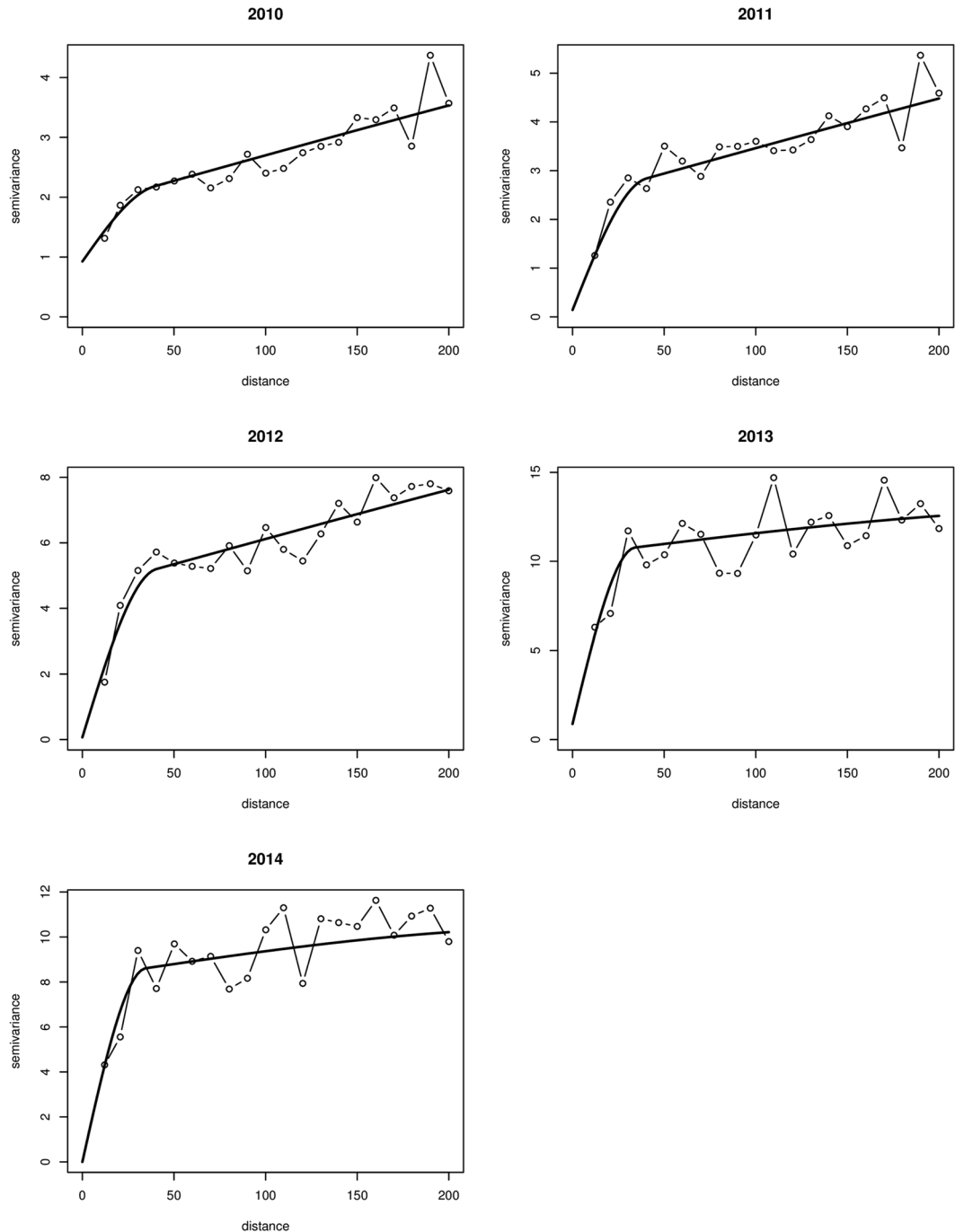


Figure 4. Nested variogram models of the risk for 2010 to 2014. This figure was created using R software (R Development Core Team 2013).

experimental variogram. All variograms exhibited two basic structures. The parameter c_0 is the nugget variance which refers to spatially random variation. The variance parameter c_1 refers to an estimate of the amount of spatially structured local (short range) variation within an average range of ϕ_1 . The variance parameter c_2 refers to an estimate of the amount of spatially structured regional (large range) variation within an average range of ϕ_2 . The sum $sill = c_0 + c_1 + c_2$ refers to the total variation, and $\%c_0, \%c_1, \%c_2$ refers to the proportion of the overall variation explained by c_0, c_1, c_2 , respectively. We found that the local spatial variations fell within the range $\approx 34\text{--}41$ km. The regional spatial variation, however, showed widespread range values, $\approx 271\text{--}900$ km, beyond the maximum lag distance used for estimation. For the years 2010 to 2012, $\hat{\gamma}_{reg}(h)$ accounted for nearly 70% or more of the variation and $\hat{\gamma}_{loc}(h)$ accounted for nearly 30% or less. Conversely, for rates in 2013 and 2014, $\hat{\gamma}_{loc}(h)$ accounted for more than 70% whereas $\hat{\gamma}_{reg}(h)$ accounted for nearly 25% or less. We also found an increasing trend of the variation accounted for by $\hat{\gamma}_{loc}(h)$ from $\approx 13\%$ in 2010 to $\approx 75\%$ in 2014.

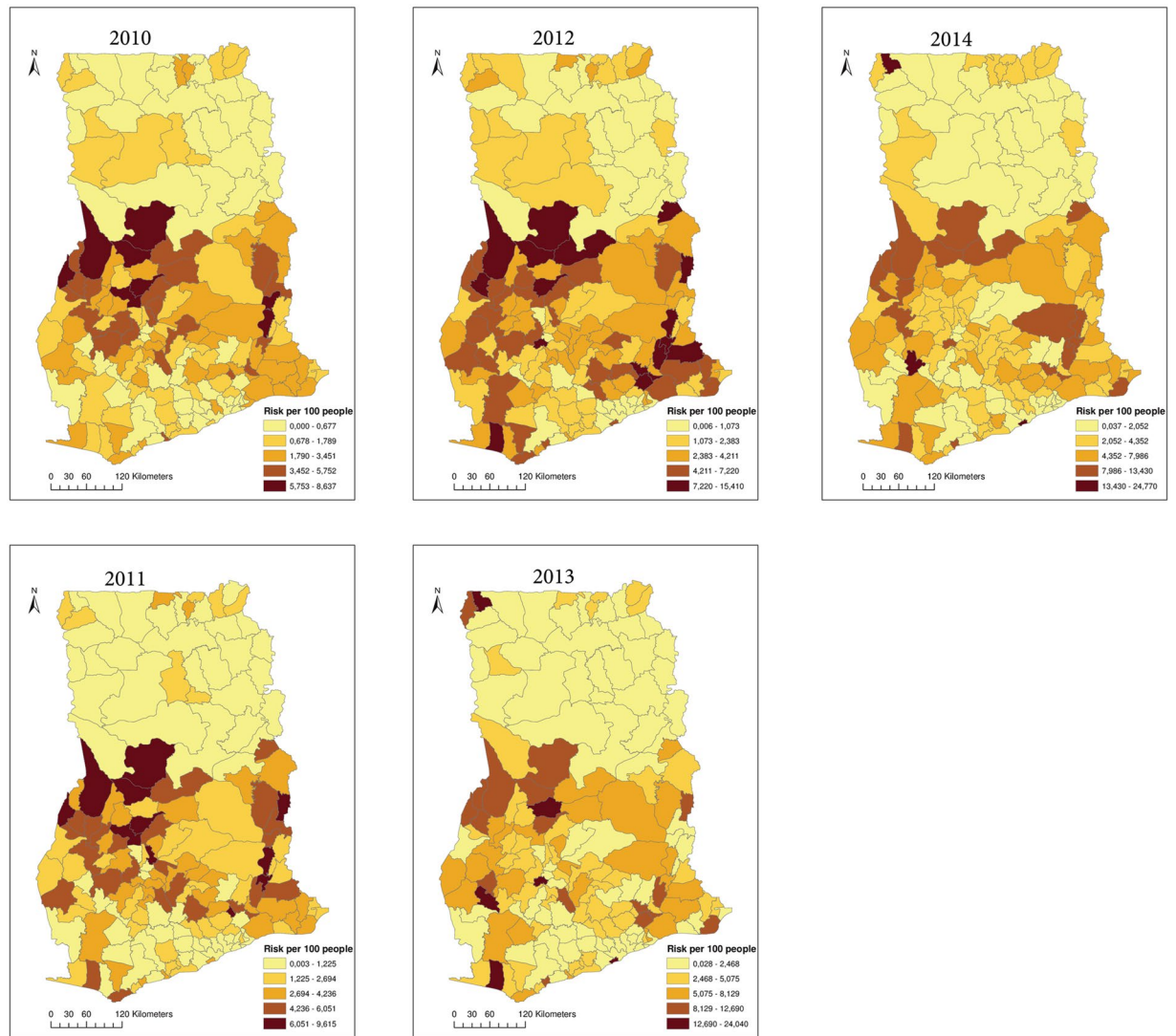


Figure 5. Maps of the smoothed rates after geostatistical filtering. This map was created using ArcGIS software (version 10.1, ESRI Inc. Redlands, CA, USA. <https://www.esri.com/>).

Year	c_0	$\%c_0$	c_1	ϕ_1 (km)	$\%c_1$	c_2	ϕ_2 (km)	$\%c_2$
2010	0.925	13.40	0.919	39.603	13.31	5.060	883.147	73.29
2011	0.141	1.62	2.277	40.399	26.14	6.293	900.055	72.24
2012	0.067	0.48	4.506	41.186	32.48	9.300	900.051	67.04
2013	0.876	6.73	9.466	34.177	72.70	2.678	314.030	20.57
2014	0.000	0.00	8.184	34.608	74.89	2.744	271.626	25.11

Table 2. Summary of the variogram models and parameters fitted to the experimental variograms.

Maps of the smoothed rates and kriging variance after geostatistical filtering are shown in Figs 5 and 6, respectively. Both the smoothed rates and the variances show clustering of districts with similar estimates. From the smoothed maps, we found that high rates dominate within the middle belt, whereas low rates dominate within the northern parts. As was expected, we found reduced variation and considerable adjustments of rates for districts with smaller populations than districts with larger populations (Fig. 7).

Discussion

We observed several noteworthy insights. We found evidence of global clustering of districts with comparable risks, suggesting the importance of spatially dependent phenomena modulating the spatial heterogeneity in the risk of intestinal worms infection. The observed global patterns imply that neighboring districts have similar

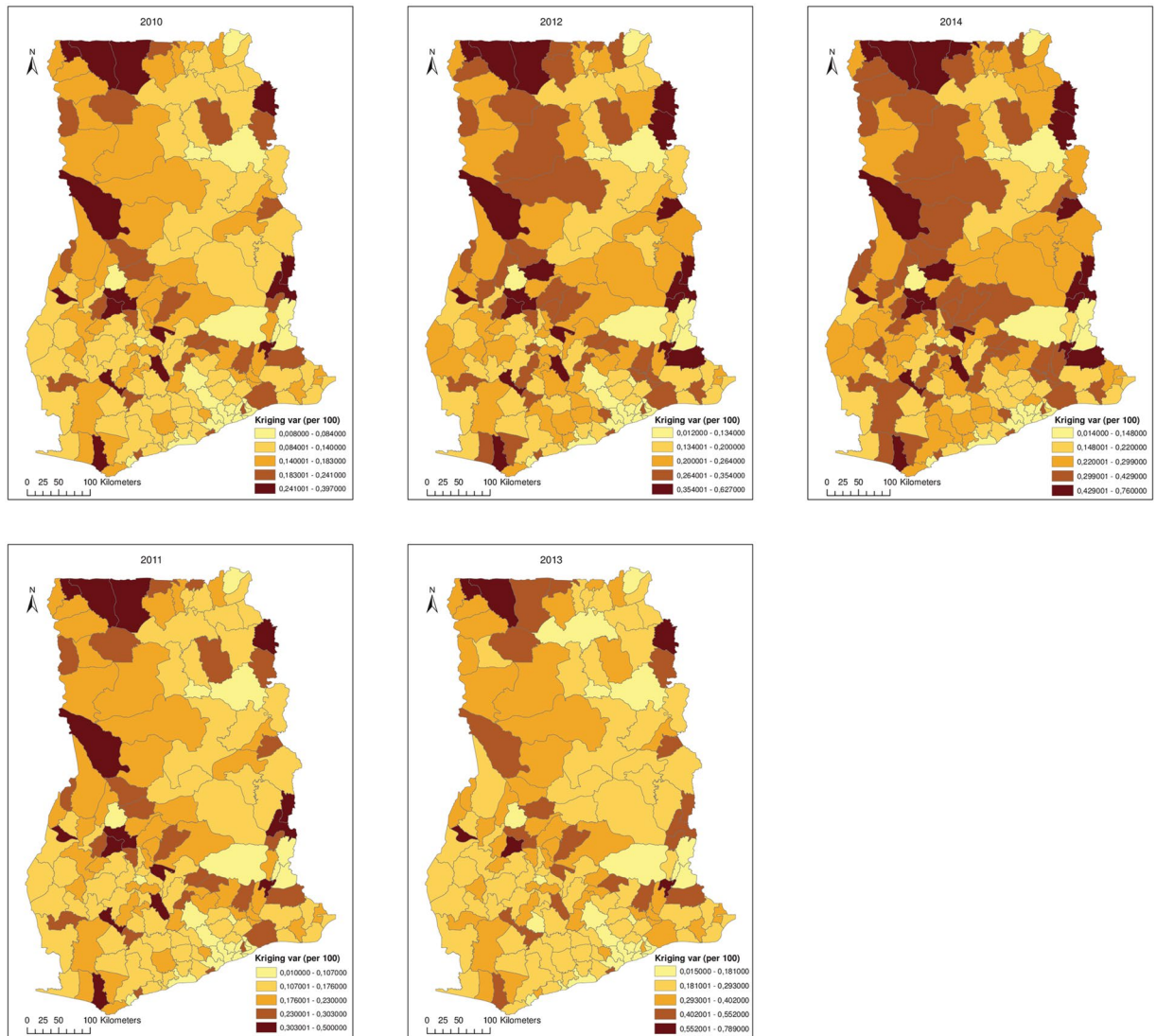


Figure 6. Maps of the Poisson kriging variance after geostatistical filtering. This map was created using ArcGIS software (version 10.1, ESRI Inc. Redlands, CA, USA. <https://www.esri.com/>).

underlying ecological and environmental risk factors that trigger intestinal worms infections²⁸. The similarity, however, widely contrasted as local Moran's *I* illuminated that the high-high values (hot-spots) dominated within the middle belt, whereas low-low values (low risk) within the northern part. One would, on the contrary, expect the hot-spots to occur mainly within the northern parts of Ghana where the socioeconomically less privilege are mostly found. This finding is an unexpected departure from other previous studies^{20,45–47}, and suggests that interaction effects of environmental and socioeconomic risk factors that combine best to enhance infection could play a role. Although the study did not build a formal causal explanation model for the clustering patterns, a visual comparison of the patterns with the ecological zones of Ghana generates a working hypothesis that is worth testing in future research efforts. The patterns of the hot-spots are plausible since they widely intersected with the semi-deciduous forest and the transitional ecological zones. High precipitation, which is mostly associated with the semi-deciduous forest and the transitional zones of Ghana, has been found to increase the risk of intestinal parasites¹⁹. The specific effect on precipitation, however, has been attributed to specific quarters of the year^{19,21}. This emphasizes the need for further studies to substantiate this augment in our study area. The low risks on the other hand widely intersect within the Guinea and Sudan Savannah ecological zones. These ecological zones are mostly flat with low precipitation, high temperature, and consist predominantly of grassland. These zones also have much drier soils with the highest land surface temperatures due to their proximity to the Sahel and the Sahara, and likely provide unfavorable environmental and ecological conditions for transmission. Some studies have associated low risk of intestinal worms infection to these unfavorable conditions. For instance²⁴, associated low risk of helminth infection with high land surface temperature in Kenya. A study of the spatial distribution of helminth infection across sub-Saharan Africa associated extreme dry soils with the absence of hookworm infections¹⁹. The patterns of hot-spots and low risks are similar to findings from a study in Cote d'Ivoire, where low risk of schistosomiasis was found in Savannah ecological zones and high risk extended into Forest ecological

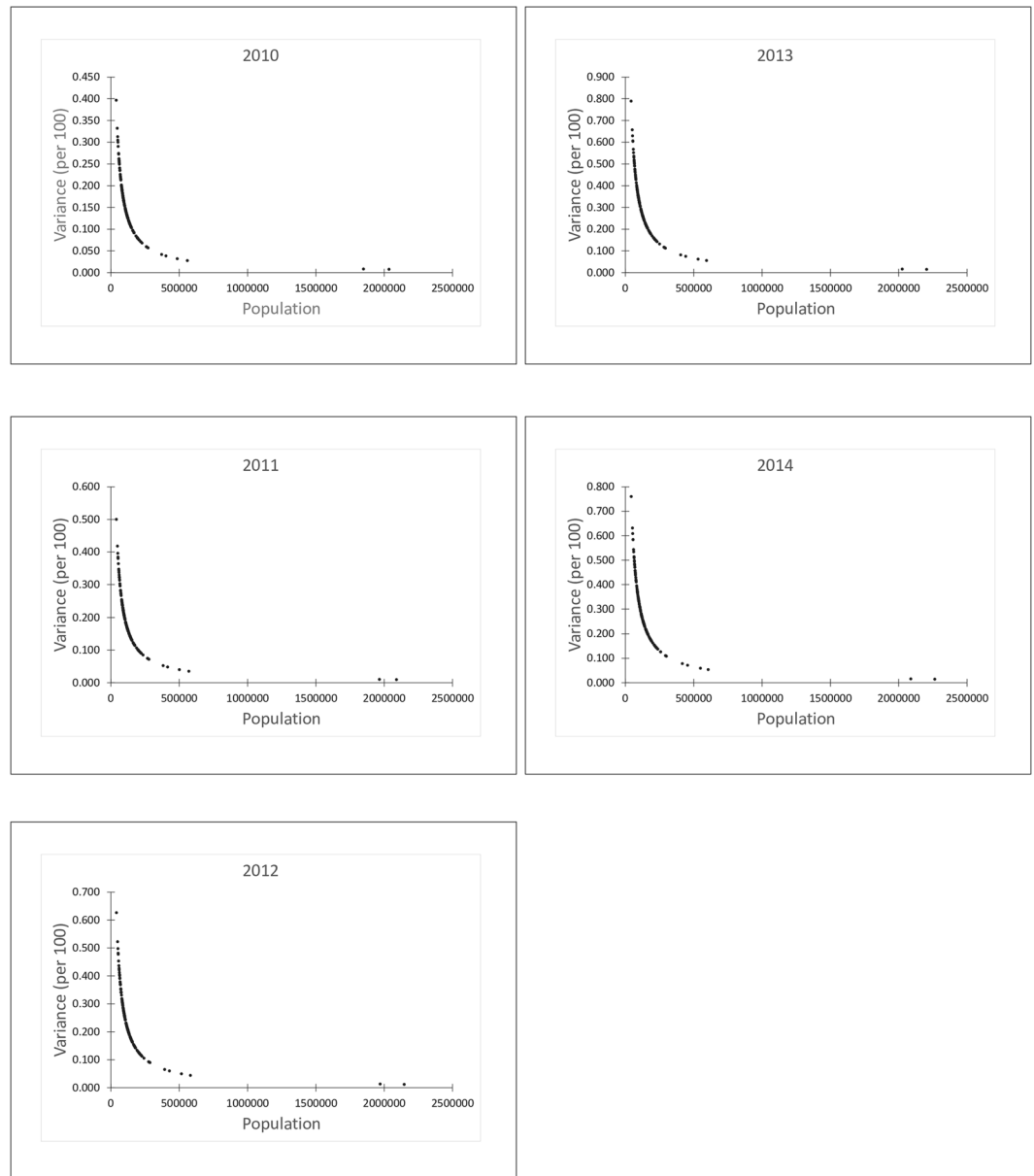


Figure 7. Graphs of variances accounted for against population after Poisson kriging.

zones⁴⁸. In a developing country like Ghana, an alternative interpretation of the clustered patterns could be based on variations in the reporting systems. However, the consistency in these patterns throughout the study period suggests that they are less likely to have been caused by only variations in the reporting systems.

The Poisson variogram estimator accounted for the effects of population heterogeneities and revealed actual spatial structures that could otherwise have been obscured from the traditional variogram estimator^{39,40}. In conjunction with the global and local statistics, nested variograms allowed us to identify the spatial distribution of intestinal parasites infection, showing that spatial variation occurs on two different scales. The average range of the local scale variation was 38 km, suggesting a strong correlation between neighboring districts. The average range of the small scale variation was less than the minimum distance of 68 km within which each district would have at least one adjoining neighbor. This suggests that most of the patterns of the small range structure only cover a district plus its adjoining neighbors. The range of the large-scale variation, $\approx 271\text{--}900$ km, was larger than the maximum lag distance probably because of the dependency on ecological processes which operate with marked variation at regional scales larger than the average size of the districts and their higher-order adjoining neighbors. The short range parameters in the second structure of the variogram in 2013 and 2014 could be due to temporal changes in the regional scale ecological processes that affect infections.

Under the assumption that the unknown risk is a spatial stochastic process, Poisson kriging of the spatial risk had the advantage to correct for extreme rates due to small populations. The Poisson kriged risk maps indicated that intestinal parasites infections are spatially varied and widely distributed. High risk markedly persisted within the middle

belt and low risks within the northern sector. This partly corroborates with data presented in a previous²³ which found low hookworm egg counts ($<1 \text{ egg g}^{-1}$) mostly in the northern parts and high hookworm egg counts ($>30 \text{ eggs g}^{-1}$) within central parts. Temporal changes in the spatial patterns over the years have been marginal, probably because the risk factors have generally remained static over time. Further studies to substantiate this argument will be valuable.

Our study still has some limitations. The first limitation relates the data. Data from the CHIM likely have recording gaps due to the voluntary reporting nature. Most reporting facilities (hospitals, clinics) lack diagnostic apparatus for proper biological confirmation of infection and hence rely on symptomatic diagnosis. However, we share the same opinion with Julian⁴⁹ that imperfect information is likely more useful to intervention design than no information. Secondly, the morbidity data covered large heterogeneous districts. Although making inferences at district level was our interest, centroid based Poisson kriging makes an implicit assumption of homogenous population and morbidity distribution within districts. This is overly simplistic and could have affected our final smoothed maps. Further studies using rigorous statistical estimations are required to attenuate possible misspecification.

Conclusions

This study demonstrated the use of spatial statistical methods such as cluster analysis and geostatistical smoothing to explore and elucidate the spatial patterns of district level intestinal parasites infections. Local and global Moran's *I* estimated and mapped spatial clustering of intestinal parasites. Our findings regarding global Moran's *I* indicated a non-random spatial distribution of internal parasites infection, and prompt for further studies to investigate and enumerate possible environmental and socioeconomic factors that could account for such patterns. Local Moran's *I* cluster maps are essential for guiding public health officials to develop cost-effective control measures and could ensure that control programs are focused appropriately. In consequence of our findings, health professionals should pay more intervention attention to the hot-spots locations. Besides, the findings regarding the pattern of global and local autocorrelations are important steps in a process leading to a proper model for intestinal worms infection in the future. Finally, the study demonstrated the usefulness of geostatistics for filtering out noise caused by heterogeneous populations, which is important for low morbidities recorded in areas with low population sites. The geostatistical risk maps provided knowledge of the spatial distribution of intestinal parasites infections Ghana. We intend to investigate issues of spatial support further and varying population and morbidity distribution within districts in the future.

References

- Cleaveland, S., Laurensen, M. K. & Taylor, L. H. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 991–999 (2001).
- Brooker, S. Estimating the global distribution and disease burden of intestinal nematode infections: Adding up the numbers – A review. *Int. J. Parasitol.* **40**, 1137–1144 (2010).
- Bethony, J. *et al.* Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *The Lancet* **367**, 1521–1532 (2006).
- de Silva, N. R. *et al.* Soil-transmitted helminth infections: updating the global picture. *Trends Parasitol.* **19**, 547–551 (2003).
- Stephenson, L. S., Latham, M. C., Adams, E. J., Kinoti, S. N. & Pertet, A. Physical fitness, growth and appetite of Kenyan school boys with hookworm, *Trichuris trichiura* and *Ascaris lumbricoides* infections are improved four months after a single dose of albendazole. *J. Nutr.* **123**, 1036–1046 (1993).
- Stephenson, L. S., Latham, M. C., Kinoti, S. N., Kurz, K. M. & Brigham, H. Improvements in physical fitness of Kenyan schoolboys infected with hookworm, *Trichuris trichiura* and *Ascaris lumbricoides* following a single dose of albendazole. *Trans. R. Soc. Trop. Med. Hyg.* **84**, 277–282 (1990).
- Stephenson, L. S., Latham, M. C., Kurz, K. M., Kinoti, S. N. & Brigham, H. Treatment with a single dose of albendazole improves growth of Kenyan schoolchildren with hookworm, *Trichuris trichiura*, and *Ascaris lumbricoides* infections. *Am. J. Trop. Med. Hyg.* **41**, 78–87 (1989).
- World Health Organization. Soil-transmitted helminthiases: STH: eliminating soil-transmitted helminthiases as a public health problem in children: progress report 2001–2010 and strategic plan 2011–2020. (World Health Organization, 2012).
- Annan, A., Crompton, D. W., Walters, D. E. & Arnold, S. E. An investigation of the prevalence of intestinal parasites in pre-school children in Ghana. *Parasitology* **92**(Pt 1), 209–217 (1986).
- Ayeh-Kumi, P. F. *et al.* Prevalence of intestinal parasitic infections among food vendors in Accra, Ghana. (2009).
- Baidoo, S. E., Tay, S. C. K. & Abruquah, H. H. Intestinal helminth infection and anaemia during pregnancy: A community based study in Ghana. *Afr. J. Microbiol. Res.* **4**, 1713–1718 (2010).
- Duedu, K., Peprah, E., Anim-Baidoo, I. & Ayeh-Kumi, P. F. Prevalence of Intestinal Parasites and Association with Malnutrition at a Ghanaian Orphanage. *Hum. Parasit. Dis.* **5**, doi:<https://doi.org/10.4137/HPD.S30059> (2015).
- Duedu, K., Karikari, Y. A., Attah, S. K. & Ayeh-Kumi, P. F. Prevalence of intestinal parasites among patients of a Ghanaian psychiatry hospital. *BMC Res. Notes* **8** (2015).
- Humphries, D. *et al.* Hookworm Infection among School Age Children in Kintampo North Municipality, Ghana: Nutritional Risk Factors and Response to Albendazole Treatment. *Am. J. Trop. Med. Hyg.* **89**, 540–548 (2013).
- Brooker, S., Clements, A. C. & Bundy, D. A. Global epidemiology, ecology and control of soil-transmitted helminth infections. *Adv. Parasitol.* **62**, 221–261 (2006).
- Brooker, S. & Michael, E. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Adv. Parasitol.* **47**, 245–288 (2000).
- Chammartin, F. *et al.* Spatio-temporal distribution of soil-transmitted helminth infections in Brazil. *Parasit. Vectors* **7**, 440 (2014).
- Fuhrmann, S. *et al.* Risk of Intestinal Parasitic Infections in People with Different Exposures to Wastewater and Fecal Sludge in Kampala, Uganda: A Cross-Sectional Study. *PLoS Negl. Trop. Dis.* **10**, e0004469 (2016).
- Karagiannis-Voules, D.-A. *et al.* Spatial and temporal distribution of soil-transmitted helminth infection in sub-Saharan Africa: a systematic review and geostatistical meta-analysis. *Lancet Infect. Dis.* **15**, 74–84 (2015).
- Nobre, L. N. *et al.* Risk factors for intestinal parasitic infections in preschoolers in a low socio-economic area, Diamantina, Brazil. *Pathog. Glob. Health* **107**, 103–106 (2013).
- Chammartin, F. *et al.* Modelling the geographical distribution of soil-transmitted helminth infections in Bolivia. *Parasit. Vectors* **6**, 152 (2013).
- Magalhães, R. J. S. *et al.* Mapping the Risk of Soil-Transmitted Helminth Infections in the Philippines. *PLoS Negl. Trop. Dis.* **9**, e0003915 (2015).
- Magalhães, R. J. S. *et al.* Mapping Helminth Co-Infection and Co-Intensity: Geostatistical Prediction in Ghana. *PLoS Negl. Trop. Dis.* **5**, e1200 (2011).

24. Pullan, R. L. *et al.* Spatial Modelling of Soil-Transmitted Helminth Infections in Kenya: A Disease Control Planning Tool. *PLoS Negl. Trop. Dis.* **5**, e958 (2011).
25. Pullan, R. L. *et al.* Human Helminth Co-Infection: Analysis of Spatial Patterns and Risk Factors in a Brazilian Community. *PLoS Negl. Trop. Dis.* **2**, e352 (2008).
26. Moran, P. A. P. Notes on continuous stochastic phenomena. **37**, 17–23 (1950).
27. Anselin, L. Local Indicators of Spatial Association—LISA. *Geogr. Anal.* **27**, 93–115 (1995).
28. Rezaeian, M., St Dunn, G., Leger, S. & Appleby, L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. *J. Epidemiol. Community Health* **61**, 98–102 (2007).
29. Beale, L., Abellan, J. J., Hodgson, S. & Jarup, L. Methodologic Issues and Approaches to Spatial Epidemiology. *Environ. Health Perspect.* **116**, 1105–1110 (2008).
30. Goovaerts, P. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *Int J Health Geogr* **4**, 31 (2005).
31. Goovaerts, P., Jacquez, G. M. & Greiling, D. Exploring scale-dependent correlations between cancer mortality rates using factorial kriging and population-weighted semivariograms. *Geogr. Anal.* **37**, 152–182 (2005).
32. Gotway, C. A. & Wolfinger, R. D. Spatial prediction of counts and rates. *Stat. Med.* **22**, 1415–1432 (2003).
33. Assunção, R. M. & Reis, E. A. A new proposal to adjust Moran's I for population density. *Stat. Med.* **18**, 2147–2162 (1999).
34. Marshall, R. J. Mapping disease and mortality rates using Empirical Bayes estimators. 40–283 (1991).
35. Bailey, T. C. & Gatrell, A. C. *Interactive spatial data analysis*. (Longman Scientific & Technical, 1995).
36. Bivand, R. & Piras, G. Comparing implementations of estimation methods for spatial econometrics. in (American Statistical Association, 2015).
37. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing (2016).
38. Mungiole, M., Pickle, L. W. & Simonson, K. H. Application of a weighted head-banging algorithm to mortality data maps. *Stat. Med.* **18**, 3201–3209 (1999).
39. Monestiez, P., Dubroca, L., Bonnin, E., Durbec, J. P. & Guinet, C. Geostatistical modeling of spatial distribution of Balanoptera phyalus in the northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts. *Ecol. Model.* **193**, 615–628 (2006).
40. Monestiez, P., Dubroca, L., Bonnin, E., Durbec, J. P. & Guinet, C. Comparison of model based geostatistical methods in ecology: application to fin whale spatial distribution in northwestern Mediterranean Sea. In *Geostatistics Banff* (eds. Leuangthong, O. & Dordrecht, D. C. V.) 777–786 (Kluwer Academic Publishers, The Netherlands, 2005).
41. Goovaerts, P., Jacquez, G. M. & Greiling, D. Exploring Scale-Dependent Correlations Between Cancer Mortality Rates Using Factorial Kriging and Population-Weighted Semivariograms. *Geogr. Anal.* **37**, 152–182 (2005).
42. Kerry, R., Goovaerts, P., Haining, R. P. & Ceccato, V. Applying Geostatistical Analysis to Crime Data: Car-Related Thefts in the Baltic States. *Geogr. Anal.* **42**, 53–77 (2010).
43. Webster, R. & Oliver, M. A. *Geostatistics for Environmental Scientists*. (John Wiley & Sons, 2007).
44. Olea, R. A. *Geostatistics for Engineers and Earth Scientists*. (Springer Science & Business Media, 2012).
45. Chin, Y. T. *et al.* Prevalence and risk factors of intestinal parasitism among two indigenous sub-ethnic groups in Peninsular Malaysia. *Infect. Dis. Poverty* **5**, 77 (2016).
46. Mehraj, V., Hatcher, J., Akhtar, S., Rafique, G. & Beg, M. A. Prevalence and Factors Associated with Intestinal Parasitic Infection among Children in an Urban Slum of Karachi. *PLoS ONE* **3** (2008).
47. Yentur Doni, N., Gürses, G., Şimşek, Z. & Yıldız Zeyrek, F. Prevalence and associated risk factors of intestinal parasites among children of farm workers in the southeastern Anatolian region of Turkey. *Ann. Agric. Environ. Med.* **22**, 438–442 (2015).
48. Chammartin, F. *et al.* Bayesian Risk Mapping and Model-Based Estimation of Schistosoma haematobium–Schistosoma mansoni Co-distribution in Côte d'Ivoire. *PLoS Negl. Trop. Dis.* **8**, e3407 (2014).
49. Julian, T. R. Environmental transmission of diarrheal pathogens in low and middle income countries. *Env. Sci. Process. Impacts* **18**, 944–955 (2016).

Acknowledgements

We extend our sincere appreciation to the Disease Control Unit of the Ghana Health Service for providing all the necessary data and background information for this research.

Author Contributions

FBO conceived of the study and carried out the analysis and drafted the manuscript. AS conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017