Spatial Variation in Risk Factors for Malaria in Muleba, Tanzania

Charles Russell Thickstun

A thesis submitted to the Faculty of Graduate and Postdoctoral

Studies in partial fulfillment of the requirements for a

Masters of Science in Epidemiology

School of Epidemiology & Public Health

Faculty of Medicine

University of Ottawa

Preface

The epidemiological and entomological data for this thesis comes from the Pan-African Malaria Vector Research Consortium (PAMVERC) cluster randomized controlled trial: *Evaluation of a novel long lasting insecticidal net and indoor residual spray product, separately and together, against malaria transmitted by pyrethroid resistant mosquitoes in Muleba, Tanzania* (PAMVERC, 2014). The principal study paper from the trial: *Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomized controlled, two-by-two factorial design trial* (Protopopoff, et al., 2018) examines the effectiveness of large-scale malaria control using alternatives to pyrethroid-only long lasting insecticidal net (LLIN) and indoor residual spray (IRS) products in an effort to identify alternative measures of malaria control in light of increasing pyrethroid resistance in vector mosquitoes.

In the course of this trial, researchers identified knowledge gaps regarding the spatial aspects of cluster-randomized control trials for malaria vector control trials in general, and malaria in particular. Specific knowledge gaps were identified regarding 1) the potential presence of a sampling bias due to standard cluster selection methods, and 2) the lack of evidence surrounding the spatial extent of cluster boundaries and buffer zones. These two issues were the initial points of inquiry from which this thesis work emerged and they were the sole concern of the author of this thesis.

Dr. Manisha Kulkarni is the Primary Supervisor of this thesis project. As a collaborator on the PAMVERC Muleba cluster-randomized controlled trial she is familiar with the data and

variables included in the dataset. Dr. Kulkarni has also served as a source of information for the spatial analytic techniques used in this project.

Dr. Ann Jolly is the Co-Supervisor of this thesis project and has provided her expertise on public health, infectious diseases, and studies involving vulnerable populations. Dr. Jolly has offered substantial advice and recommendations for ensuring this thesis is clear and comprehensible to readers from a variety of backgrounds.

Dr. Alice Zwerling has provided additional resources on epidemiological and statistical methods, model construction, global health settings, and the health economic and programmatic implications of this project.

The dataset for both parts of this thesis project have been provided by Dr. Manisha Kulkarni, the primary supervisory of this project, and Dr. Natacha Protopopoff, a principal investigator on the PAMVERC trial. Research ethics approval for secondary analysis of the data obtained in the PAMVERC trial has been granted to Dr. Kulkarni by the Ottawa Health Science Network Research Ethics Board (Protocol # 20150199-01H). Approval to add the author of this thesis as an additional researcher on the project was granted on 30 October, 2017. Initial approval was granted until 2 June, 2018 and was subsequently extended for a further year to 1 June, 2019. Additional ethics approval has been granted by the University of Ottawa Science and Health Sciences Research Ethics Board for the secondary analysis of a pre-existing dataset pertaining to this thesis through 31 October, 2019 (Ethics File Number H-06-18-806). This thesis project is, in its entirety, a secondary data analysis and as such no primary data collection has been conducted.

Glossary of Key Terms

Actellic®300CS – An Indoor Residual Spray (IRS) product which uses a micro-encapsulated pirimiphos-methyl organophosphate insecticidal agent instead of standard permethrin. Used as novel intervention in the PAMVERC Muleba trial.

Anopheles species – (*Anopheles spp.)* Mosquito vectors most responsible for malaria infection in the study area. Predominately *An. gambiae* sensu stricto (s.s.), *An. arabiensis* and *An. funestus* s.s.

Cluster (cRCT) – Grouping of households by geographic or political boundaries to facilitate allocation and distribution of interventions in a cluster randomized controlled trial (cRCT). Often represented by a Core sampling area and a Buffer where samples are not collected to prevent contamination.

Cluster (Spatial Statistics) – A higher/lower than expected concentration of events (or values of events) in a geographical area. May be local (cluster detection), or global (sometimes referred to as "clustering" or spatial autocorrelation).

Community Effect – An effect wherein individuals who are living in a community with sufficient intervention coverage benefit from the protective effect even if they do no partake in the intervention. Results from a mass reduction in mosquito vector population density and mean age.

Contamination or Spillover Effect – A result of community effects being shared between proximal intervention clusters from different allocations. Results in the difference between groups skewing towards the null hypothesis.

Global and Local Effects – Effects from a statistical process that are global represent the total effect across the study area. Local effects take spatial proximity into consideration by restricting or weighting inputs based on distance values.

Indoor Residual Spraying (IRS) – Insecticide spray formulation designed to remain on the interior walls of a household and persistently kill or repel mosquito vectors.

Insecticide Treated Net (ITN) – Bednet that has been soaked in an insecticide solution. Needs to be retreated regularly to remain effective.

Long-lasting Insecticidal Net (LLIN) – Bednet that has an insecticide bound to the net fibres during the manufacturing process. Remains effective for a number of years.

Malaria – Acute febrile illness caused by infection with *Plasmodium* species parasites. Has high morbidity and mortality if not treated rapidly in susceptible hosts. Infection may be present without clinical symptoms in hosts with strong immune response.

Olyset – Standard WHO recommended single insecticide (permethrin, a form of pyrethroid insecticide) treated bednet. Used as control in PAMVERC Muleba study.

Olyset Plus – Single insecticide (permethrin) treated bednet, with an added formulation (Piperonyl Butoxide) which inhibits insecticide resistance in mosquito vectors. Novel intervention in PAMVERC Muleba study.

Plasmodium species – (*Plasmodium spp.*) The parasitic agent that causes malaria. Infection may be asymptomatic or cause acute fever. Predominately *P. falciparum* in the study area.

Pyrethroid – A class of insecticide with exito-repellent and knockdown (i.e. paralysis) effects on mosquito species.

Semivariance – One half the mean squared difference in a set of measurements from two locations. May be calculated with more robust statistics using M-estimation.

Semivariogram – A model of spatial autocorrelation represented by semivariance as a function of distance. Characterized by the Sill (limit of the fitted semivariogram function), Range (distance at which the model reaches the Sill), Nugget (discontinuity in semivariance at the origin), and Partial Sill (scale of the semivariance in the study area, minus the nugget discontinuity).

Spatial Autocorrelation – Tendency of measurements to be more similar at smaller distances. Often calculated with a Moran's I statistic or described by computing a semivariogram.

Stationarity – Homogeneity of a process across a study area. In opposition to non-stationarity, in which the process repeats itself irregularly. Quasi-stationarity exists when a non-stationary process is stationary within the sample window.

Vector – An intermediary (usually an arthropod) that transports an infectious agent to a host.

Abstract

Despite the rich knowledge surrounding risk factors for malaria, the spatial processes of malaria transmission and vector control interventions are underexplored. This thesis aims 1) to describe the spatial variation of risk factor effects on malaria infection, and 2) to determine the presence and range of any community effect from malaria vector control interventions. Data from a cluster-randomized control trial in Tanzania were analyzed to determine the geographically-weighted odds of malaria infection in children at trial baseline and post-intervention. The spatial range of intervention effects on malaria infection was estimated post-intervention using semivariance models. Spatial heterogeneities in malaria infection and each covariate under study were found. The median effective semivariance range of intervention effects was approximately 1200 meters, suggesting the presence of a community effect that may cause contamination between trial clusters. Trials should consider these spatial effects when examining interventions and ensure that clusters are adequately insulated from contamination.

Keywords Malaria, Geographically-weighted Regression, Cluster Detection, Community Effect, Buffer, Semivariance, Spatial Analysis

Acknowledgements

I would like to begin by acknowledging that this thesis was conceived, analyzed, and written on the traditional and unceded territory of the Algonquin nation. While this thesis does not directly investigate matters surrounding Canadian indigenous communities, it does involve peoples and cultures distinct from that of the author that must be similarly respected. Every effort has been made to ensure that local cultural norms and practices have been built into the assumptions made in this analysis, and that any information gleaned from this study will be used to the benefit of Tanzania and its people.

I would like to thank Dr. Manisha Kulkarni for providing me with the opportunity to work on this project, for bearing with me through an enormous number of revisions and rewrites, and for guiding me through the complications and processes of being a young scholar entering the field of Global Health. I would like to thank Dr. Ann Jolly for her help and encouragement in many respects over the three years I have been in the program, and I would like to thank Dr. Alice Zwerling for her feedback throughout the process, and her reassurance that I would be able to achieve the goals I had set. My Thesis Advisory Committee has been a consistent support during this project, and for that I cannot thank them enough.

I would like to thank the PAMVERC team for their enormous efforts to collect and collate this data in order to enable the analyses I have conducted. I would also like to thank the members of the PAMVERC Misungwi Team for giving me the opportunity to present the preliminary analysis of community effect ranges at the project Kick-off meeting. I would like to especially thank Dr. Natacha Protopopoff for her continued interest in the project and its implications.

I would like to thank Dr. William Stanford for chairing my defense, and Drs. Yue Chen and Daniel Corsi for agreeing to be examiners for my thesis, and for the valuable feedback they have provided in the process.

I would like to thank everyone in the INSIGHT lab for their help and feedback on many iterations of this thesis and its analyses. Thank you to Isha and Mariame for guiding me through my first year in the program, and thank you to Jessica, Andreea, Jay, Roman, Benoit, and Aisha for sitting through far too many presentations.

And finally, I would like to thank my friends and family for providing continued support. Mom, Dad, John, Taryn, Lauren, I couldn't have made it through without you.

Table of Contents

Table of Figures

Table of Tables

Acronyms

95%CI – Ninety-Five Percent Confidence Interval

ACT – Artemisinin-based Combination Therapy

AICc – Alkaline Information Criteria (corrected)

AIDS – Acquired Immune Deficiency Syndrome

ANOVA – Analysis of Variance

cRCT – Cluster Randomized-Control Trial

DALY – Disability-Adjusted Life Year

GWLR – Geographically Weighted Logistic Regression Analysis

GWR – Geographically Weighted Regression Analysis

HIV – Human Immunodeficiency Virus

IRS – Indoor Residual Spraying

ITN – Insecticide Treated Net

IVCC – Innovative Vector Control Consortium

KCMCo – Kilimanjaro Christian Medical College

LLIN – Long-lasting Insecticidal Net

LSHTM – London School of Hygiene and Tropical Medicine

LSM – Larval Source Management

MI – Moran's I Statistic

NIMR – National Institute for Medical Research (Tanzania)

O/E – Observed/Expected

OLS – Ordinary Least Squares

PAMVERC – Pan-African Malaria Vector Research Consortium

PBO – Piperonyl Butoxide

RCT – Randomized-Control Trial

RDT – Rapid Diagnostic Test

TDR – Special Programme for Research & Training in Tropical Diseases (WHO)

VCAG – Vector Control Advisory Group (WHO)

WGS_1984 – World Geodetic System 1984

WLS – Weighted Least Squares

Software

ArcGIS – Mapping and Analytics Platform (ArcGIS, 2015)

GWR4 – Geographically Weighted Modelling (GWR4, 2015)

SAS – Analytics, Business Intelligence and Data Management (SAS, 2012)

SaTScan – Software for the spatial, temporal, and space-time scan statistics (SaTScan, 2016)

1. Introduction

It is easy to note that a communicable disease is spread by exposure to its agent of infection and suggest that one must simply avoid that particular bacteria, virus, or parasite to remain healthy. It may become harder to maintain that confidence once it is discovered that the agent of infection is carried by a vector species so resistant to human intervention that efforts to remove it entirely from populous areas have failed even in the wealthiest nations. And that confidence may all but erode on the discovery that the first global program for the eradication of malaria was adopted by the Eighth World Health Assembly in 1955 (WHA, 1955), and yet the only substantial progress towards that goal has come in the last few decades.

While the underlying agent for malaria is simple enough to discern with modern techniques, assessing the risk factors associated with malaria is an ongoing process. Identifying and understanding these risk factors is essential for the development of disease prevention and control programs and lifestyle changes that could have meaningful impact. National and international level surveillance can begin to paint some of the picture, but in-depth studies must be used to evaluate population-specific risk factors that could otherwise be missed.

Using data from a cluster Randomized-Control Trial (cRCT) in the Kagera region of Tanzania, this thesis seeks to address certain gaps in knowledge of local-scale malaria transmission by 1) adding evidence to the nature of risk factors associated with human malaria infection and describing how those factors may vary spatially across a region, and 2) determining the presence of community effects of malaria vector control interventions on household malaria infection, and examining the implications these effects could have on the significance of trial outcomes.

1.1 Malaria

Malaria is an acute febrile disease caused by *Plasmodium species* parasites that are transmitted by the bite of an *Anopheles species* mosquito (APHA, 2015). Treatment with a course of antimalarial medication must begin immediately once febrile symptoms occur, as certain strains of *Plasmodium spp.,* in particular *P. falciparum*, can cause severe complications rapidly (APHA, 2015). Artemisinin-based combination therapy (ACT) is currently recommended as the first line treatment in most malaria cases with a confirmed diagnosis (WHO, 2015). Children and pregnant women are the most vulnerable to malaria as well as those who are immunocompromised (CDC, 2016). The case-fatality rate for these groups can reach 20% once complications or severe malaria develops (APHA, 2015).

Malaria is a high-burden disease globally, with ongoing transmission in 91 countries, and contributing to an estimated 219 million cases and 435,000 deaths in 2017 (WHO, 2018). Ninety-two percent of these cases and 93% of deaths from malaria occurred in the World Health Organization (WHO) African Region (WHO, 2018) where *P. falciparum* predominates (Gething, et al., 2011) with *Anopheles gambiae* and *An. funestus* as primary vectors (Sinka, et al., 2012; Wiebe, et al., 2017). The United Nations has set malaria as a priority for control with Millennium Development Goal 6 targeting the reduction of cases by 2015 (United Nations, 2000), and Sustainable Development Goal 3.3.3 targeting the elimination of malaria by 2030 (United Nations, 2015). Global Disease Burden for malaria has decreased substantially since the initial targets were set in 2000, reaching its peak of 3.55% of total Disability-Adjusted Life Years (DALYs) in 2005 and decreasing to 2.27% by 2015 (IHME, 2016). Cases of malaria have decreased during this period, from a high of 238 million in 2000 to a low of 214.2 million in 2015, but have since seen a resurgence in the last two years causing concerns with current

control practices (WHO, 2018). Deaths from malaria have seen no similar resurgence and continue to decline worldwide (WHO, 2018).

1.2 Risk Factors for Malaria

Factors associated with malaria infection or clinical disease have been explored extensively. These act as a combination of factors influencing the *Plasmodium* parasite, the *Anopheles* vector, and the human host they come in contact with (Protopopoff, et al., 2009). Many studies have focused on identifying modifiable risk factors for malaria, i.e. those that can either be targeted for intervention or changed to provide protection from malaria, and can be assessed without enormous cost. In recent years major malaria studies have found associations with demographic variables like age, socioeconomic status, housing quality, population density, and head of household education (Walldorf, et al., 2015; Roberts & Matthews, 2016; Bannister-Tyrrell, et al., 2018; Tonye, Kouambeng, Wounang, & Vounatsou, 2018; Were, et al., 2018; Homan, et al., 2016) and environmental factors like precipitation, elevation, and temperature (Hasyim, et al., 2018; Thomson, Mason, Phindela, & Connor, 2005; Sena, Deressa, & Ali, 2015; Walldorf, et al., 2015; Daygena, Massebo, & Lindtjørn, 2017; Midekisa, Beyene, Mihretie, Bayabil, & Wimberly, 2015; Sena, Deressa, & Ali, 2015).

1.3 Prevention and Control Measures

The WHO currently recommends a combination of Pyrethroid-Only Long-Lasting Insecticidal Nets (LLINs) and Indoor Residual Spraying (IRS) with one of three approved single

insecticide formulations as the core intervention for malaria control (WHO, 2017).[1] In extreme circumstances the WHO also recommends Larval Source Management (LSM) in order to reduce vector populations (WHO, 2017). These recommendations have been maintained with reasonable consistency as far back as 2005 (WHO, 2005), although the emphasis has shifted from favouring IRS to favouring LLINs over this period. Numerous earlier studies on the efficacy of insecticide-treated net (ITN) interventions (Snow, Rowan, Lindsay, & Greenwood, 1988; Lindsay, et al., 1989; Alonso, et al., 1991; Kamol-Ratanakul & Prasittisuk, 1992; Luxemburger, et al., 1994; D'Alessandro, et al., 1995; Kroeger, 1995; Binka, et al., 1998; Nevil, et al., 1996; Zaim, Ghavami, Nazari, Edrissian, & Nateghpour, 1998) as well as the 2009 Cochrane Review on ITNs have helped cement these recommendations into practice (Lengeler, 2009). Aside from providing a physical barrier to malaria vectors which bite at night, ITNs and LLINs both use insecticide formulations to kill mosquitoes that make contact with the net. Insecticides are incorporated into the net either by periodically re-treating the entire net with insecticide (ITNs) or binding the insecticide to the fibres during the manufacturing process (LLINs) (WHO, 2007). Pyrethroid insecticides in particular have strong exito-repellent and knockdown (i.e. paralysis) effects on major malaria vectors adding to their efficacy (Kawada, et al., 2014). At sufficient coverage[2], this combination of effects leads to a reduction of mosquito populations in the area (Binka, et al., 1998; Hawley, et al., 2003) and a reduction in mean mosquito age which inhibits the ability of the mosquito to transmit sporogonic-stage parasites (Curtis, Maxwell, Magesa, Rwegoshora, & Wilkes, 2006; Govella, Okumu, & Killeen, 2010). This results in what is known as a

[1] The WHO issued an interim endorsement of nets treated with a pyrethroid and piperonyl butoxide (PBO) in late 2017 based on the creation of a new class of recommendations based on the promising results a single study. Full WHO recommendations require at least two epidemiological studies (WHO, 2017).

[2] While 35-65% total population coverage seems to be sufficient for this community effect (WHO, 2007; Killeen, et al., 2007), current recommendations call for universal coverage, defined as 80% (WHO, 2017).

"community effect", wherein individuals who are living in a community with high LLIN coverage benefit from the protective effect even if they do not sleep under a LLIN themselves (WHO, 2018).

The 2016 World Malaria Report indicates that there has been a significant reduction in coverage of population-at-risk in IRS, especially in the WHO Africa Region (from 5.7% population at risk protected by IRS in 2010 to 3.1% in 2015), corresponding to an increase in pyrethroid resistance among the main malaria vector mosquito species (WHO, 2016). This shift away from IRS is likely due to its inefficacy in the presence of pyrethroid resistance as the 2017 World Malaria Report notes that the proportion of countries with endemic malaria transmission that both monitored and reported resistance to pyrethroid insecticides increased from 71% in 2010 to 81% in 2016 (WHO, 2017).

A multi-country evaluation conducted by the WHO found that during this same period, pyrethroid-based LLINs remained effective in areas with pyrethroid resistance, although spraying with a non-pyrethroid IRS nearly halved malaria incidence (WHO, 2016).

1.4 Malaria in Tanzania

Malaria is a serious concern in Tanzania, with the country contributing an estimated 6,477,825 malaria cases in 2017 (WHO, 2018), an increase of almost 400,000 cases from 2015 (WHO, 2018). Malaria accounts for approximately 5% of total disease burden (DALYs) in Tanzania as of 2015 (IHME, 2016), and Kagera region where Muleba district is located has a malaria prevalence of 8.3% in children ages 6-59 months (PMI, 2016). The National Malaria Control Programme has outlined a Malaria Strategic Plan for Tanzania focusing on improving vector control measures, prompt diagnosis and treatment, public empowerment, information

dissemination, and financial management of malaria programs (NMCP, 2014). A massive influx of 29.2 million ITNs were delivered to Tanzania between 2014-2016 in an attempt to bring malaria vectors under control, but the modelled proportion of population with access to ITNs remained below 80% population coverage targets at an estimated 50%, with household surveys placing the percentage of households with access to enough bednets for household population at only 38.8% (WHO, 2018).

The 2016 report from the President's Malaria Initiative paints a relative contrast to this, showing strong improvements towards the goals outlined in the National Malaria Control Programme. It states that the number of households with an ITN has gone from 23% in 2004 to 91% in 2012, and prevalence of malaria parasitemia in children 6-59 months has halved from 18.1% to 9% between 2007 and 2011 (PMI, 2016), with increases in bednet use lagging behind ITN ownership in children under 5 (16% - 73%) and pregnant women (15% - 76%) during the same period (PMI, 2016). These increases are encouragingly close to the 80% target for universal coverage of vector control set by the WHO (WHO, 2017), but response time to illness, has remained low, with only 14% of children under five receiving ACT within 24 hours of symptom onset in 2008, rising to 27% in 2010, but then falling off to 21% in 2012 (PMI, 2016).

1.5 Thesis Outline

Due to the persistent burden of malaria in sub-Sahara Africa and the rising threat of pyrethroid resistance which may undermine recent progress in malaria control, a pipeline of new vector control products is under development (IVCC, 2018). These new products aim to mitigate the development of resistance by combining a pyrethroid insecticide with a synergist or other class of insecticide with different mode of action (Jaffer, et al., 2015; Protopopoff, et al., 2018).

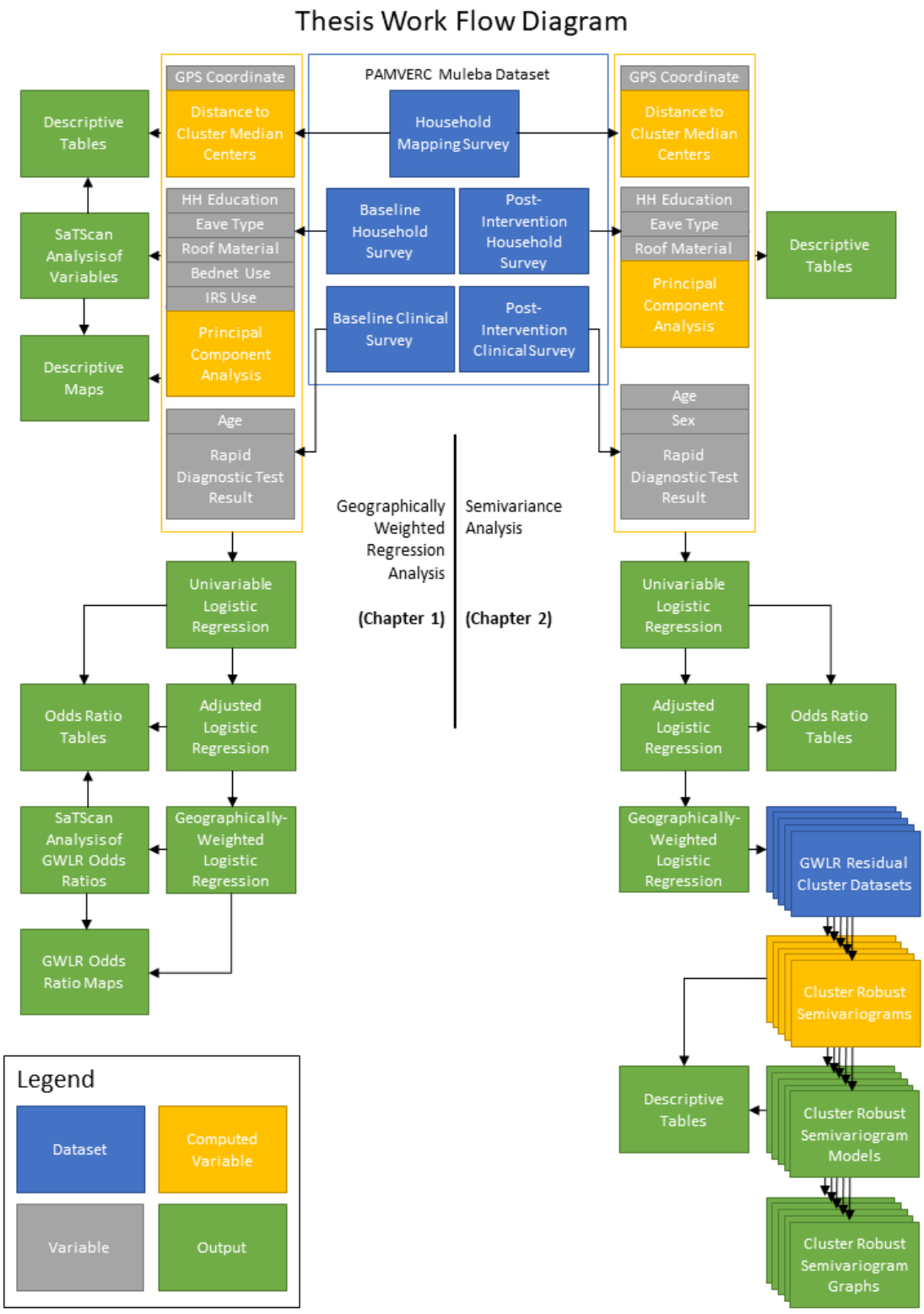
New products are evaluated in laboratory (Phase I) and semi-field (Phase II) settings, including Phase III evaluations in community settings using large-scale cluster randomized controlled trials (VCAG, 2017). However, despite the rich knowledge base on risk factors for malaria in different regions and populations, knowledge on the underlying spatial processes affecting malaria transmission at the population level, and the spatial effects of malaria vector control interventions in community settings, is limited.

In this thesis I seek to address key knowledge gaps regarding the spatial aspects of malaria transmission and malaria vector control intervention impact at a local scale. Specifically, I aim to 1) describe the spatial variation of risk factor effects on malaria infection; and to 2) determine the presence and range of a community effect on household malaria infection measured in trial clusters four months post-intervention, in children aged six months to 14 years old in Muleba, Tanzania. By establishing the spatially trending covariates related to malaria infection, it should be possible to build these into a geographically-weighted logistic regression model and capture the underlying community effect in the extracted residual effects, and then model their forms with semivariance analysis.

I begin my thesis with a review of the literature on malaria risk factors at the individual, household, and community levels and present relevant background on cRCTs for malaria vector control interventions. Following this, I describe the methodological aspects of my thesis including a complete descriptive discussion of the study dataset used, the construction of variables used in the following analyses, and a complete analytic process for each of my articles. A Work Flow Diagram (***Figure 1.1***) provides a map of which data were used for each paper in this study, and provides insight into how the inferences from the first paper were used to influence decisions made in the second. My first paper, *By All Odds: A Geographically-*

Weighted Regression Analysis of Risk Factors for Malaria in Muleba, Tanzania, addresses my first thesis objective. It identifies key factors that contribute to the prevalence of malaria infection in children pre-intervention and assesses whether these factors trend geographically across the study area. My second paper, *Inter-Cluster Contamination: A Semivariance Analysis of Community Effect Ranges in a Four-Armed Malaria Trial in Muleba, Tanzania* builds on this initial analysis to determine the range effect of LLIN and IRS interventions, adjusting for relevant factors identified in paper 1 as being both significantly associated and spatially trending. Finally, I conclude with a discussion of the results of each of my chapter analyses in context with the current state of the literature, and how they may impact research around cluster-randomized malaria trials in the future.

Figure 1.1 *Thesis Work Flow Diagram*



2. Literature Review

2.1 Risk Factors Associated with Malaria

While it has long been established that the transmission of malaria species parasites to humans occurs through the bite of an infected vector mosquito (Manson, 1900; Grassi, Bignami, & Bastianelli, 1899), numerous factors can be used to indicate the likelihood of this occurrence. While the term "risk" inherently implies an increase in the chances of malaria associated with a given factor, this thesis considers risk factors on a continuum from aggravating to protective, as the direction of association may change between – and indeed within – studies, as will be explored later in this thesis. In modern malaria studies, risk factors are usually considered at the individual level, household level, and, more recently, the community level.

2.1.1. Individual Level Risk Factors

As infectious diseases are ultimately the result of an infectious agent penetrating the immune defenses of an individual, the primary risk factors for acquiring any infectious disease must include the infected host. While it would be exceptionally resource intensive to examine the immune susceptibility and exposure profile of each individual in the study area, there are individual-level variables that can be considered as proxies to establish some measure of control in the analysis.

Age has been consistently shown to be a risk factor for malaria, with children under 15 displaying the greatest risk of malaria morbidity and mortality (West, et al., 2013; Roberts & Matthews, 2016; Were, et al., 2018; Bannister-Tyrrell, et al., 2018; Tonye, Kouambeng, Wounang, & Vounatsou, 2018; Walldorf, et al., 2015; Mwesigwa, et al., 2017). Prevalence of *P.*

falciparum infection has been shown to increase with age (Were, et al., 2018; Froeschl, et al., 2018; Bannister-Tyrrell, et al., 2018), while malaria cases (denoted by an associated acute fever) tend to be most common in children under 5 (Were, et al., 2018). This discrepancy is attributable to an increasing cumulative exposure to malaria parasites with age, while building immunity to the symptoms of the disease (Marsh, 1992; Bull, et al., 1998).

While included in most studies as a matter of form, sex has not typically shown an association with either malaria cases or malaria infection (Froeschl, et al., 2018; Maziarz, et al., 2018; Bannister-Tyrrell, et al., 2018; Tonye, Kouambeng, Wounang, & Vounatsou, 2018; Ghebreyesus, et al., 2000). Initial points of inquiry for including sex as a variable are built on the assumption that there may be gendered differences between groups, rather than sex-related differences in immune responses, such as outdoor activities or preferential use of preventive measures which could introduce more potential for contact with a malaria infected mosquito. Where it does show a positive association, boys are generally at higher risk of malaria (Brooker, et al., 2004; Mwesigwa, et al., 2017).

2.1.2. Household Level Risk Factors

Several household-level factors, including socioeconomic status (SES), housing quality, and head of household education have been associated with malaria infection in many studies. Socioeconomic status, when defined as ownership of durable physical assets, has been repeatedly shown to be a strong predictor of malaria infection and cases (Protopopoff, et al., 2009; West, et al., 2013; Homenauth E. , 2016; Were, et al., 2018; Froeschl, et al., 2018; Tonye, Kouambeng, Wounang, & Vounatsou, 2018). Like with most diseases it shares an association with, poverty exacerbates the conditions and circumstances that allow malaria to thrive. Due to the fact that

many households in the developing world do not have a reportable means of income it has become common to use principal component analysis (PCA) to derive a first-component variance of durable assets, household construction, and access to utilities as a means of building a wealth index in a study population (Vyas & Kumaranayake, 2006). These indices have been used to estimate SES in a multitude of studies and reports, either for directly observing SES associations with malaria (Homenauth, Kajeguka, & Kulkarni, 2017; Brooker, et al., 2004), or as a control variable in other studies on the depreciation of LLIN efficacy over time (Solomon, et al., 2018), or assessing climate impacts on malaria transmission (Upadhyayula, Mutheneni, Chenna, Parasaram, & Kadiri, 2015). It is the current standard measure for computing wealth indices for the Demographic and Housing Survey of Tanzania (MoHCDGEC, Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS), & ICF, 2016).

Various measures of housing quality have shown associations with malaria including flooring, wall, and roofing materials (Protopopoff, et al., 2009; Homenauth E. , 2016; Roberts & Matthews, 2016; Homenauth, Kajeguka, & Kulkarni, 2017; Walldorf, et al., 2015; Ghebreyesus, et al., 2000), and whether the house is built with open eaves – a gap between the walls of the house and the roof (Walldorf, et al., 2015; Bradley, et al., 2013; Ghebreyesus, et al., 2000). These housing quality factors are most likely associated with malaria through the mediating factor of the *Anopheles* mosquito that delivers the *Plasmodium* parasite, as they provide a mechanism of access to the host that would not be available otherwise. While these household attributes are generally correlated with socio-economic status, the above studies all examine them in models with a measure of SES control and still find them significant. Some studies have attempted to include variables in the SES indices while also including them as covariates in

statistical models to mixed success (Houweling, Kunst, & Mackenbach, 2003), but in general it seems prudent to ensure that variables to be included in the analysis be excluded from the SES index (Howe, Hargreaves, & Huttly, 2008).

The education level of an authority figure has shown an association with childhood malaria in multiple studies, whether the head of household (or spouse)[3] as was the case in a study in Malawi (Walldorf, et al., 2015), a "caregiver" as in a study in Uganda (Roberts & Matthews, 2016), or the mother as was the case in a study in Cameroon (Tonye, Kouambeng, Wounang, & Vounatsou, 2018). In all of these cases higher levels of education were consistently associated with lower risk of malaria for the children under study. Head of household education has also been associated with ITN ownership and use (Atieli, et al., 2011). Knowledge of the modes of transmission (Gessler, et al., 1995; Mboera, et al., 2010) and preventive measures and their applications (Winch, et al., 1997) can provide enormous benefits to malaria control. A systematic annotated bibliography of the socio-economic aspects of malaria, and how knowledge of malaria risk factors can contribute to reductions in malaria cases, is covered in a 2003 report by the Special Programme for Research & Training in Tropical Diseases (TDR) (Heggenhougen, Hackethal, & Vivek, 2003).

2.1.3. Community-Level Risk Factors

There are significant community-level effects that can impact malaria prevalence (Binka, et al., 1998; Howard, et al., 2000; Hawley, et al., 2003). Population density has been shown to

[3] Head of households are predominately male in the Kagera Region although are not universally so at 68.1% according to the 2012 census (National Bureau of Statistics, 2016). Additionally, the Muleba baseline household survey did not ask for maternal education, and the post-intervention survey had substantial missing data on this variable. As such this thesis uses Head of household education as a covariate, as defined by each household at the time of interview.

have a significant effect on malaria prevalence (Homan, et al., 2016; Tonye, Kouambeng, Wounang, & Vounatsou, 2018), as well as community ITN ownership (West, et al., 2013), community IRS programs (Maziarz, et al., 2018), community SES (West, et al., 2013), and malaria transmission intensity (West, et al., 2013; Homenauth E. , 2016).

Since malaria transmission intensity is driven by the density of and prevalence of infection within vector populations, any factors that promote *Anopheles spp.* mosquito breeding and reproduction are also important to consider. The primary vector species in Tanzania belong to the *An. gambiae* and *An. funestus* species complexes, with *An. gambiae* sensu stricto (s.s.), *An. arabiensis* and *An. funestus* s.s. comprising the most important vectors. Other common mosquito species, such as *Aedes aegypti* and *Culex quinquefasciatus*, are considered nuisance-biting mosquitoes and may be vectors of other pathogens such as arboviruses and lymphatic filariasis. Mosquito vector populations have been shown to have associations with precipitation (Protopopoff, et al., 2009; Ezanno, et al., 2015; Pakdad, et al., 2017), elevation (Protopopoff, et al., 2009), and temperature (Protopopoff, et al., 2009; Ezanno, et al., 2015; Asare, Tompkins, Amekudzi, Ermert, & Redl, 2016; Pakdad, et al., 2017), owing to their aquatic breeding habitat. Temperature increases have been associated with *Anopheline* survival rates (Bayoh & Lindsay, 2004; Christiansen-Jucht, Parham, Saddler, Koella, & Basáñez, 2014; Kristan, Abeku, & Lines, 2018; Lyons, Coetzee, & Chown, 2013) with a peak of survival around 25-30C (Lymio, Takken, & Koella, 1992). This is also tied to age of pupation (Lymio, Takken, & Koella, 1992; Bayoh & Lindsay, 2003; Lyons, Coetzee, & Chown, 2013), and transmission rates (Kelly-Hope, Hemingway, & McKenzie, 2009; Mordecai, et al., 2013). Additionally, temperature has been shown to impact the efficacy of some larvicidal agents on *Aedes, Anopheles,* and *Culex spp.* (Kajla, et al., 2016) which may further exacerbate the correlation, as well as parasite

development (Paaijmans, Blanford, Chen, & Thomas, 2012; Blanford, et al., 2013), which may impact the association with malaria. *Anopheles spp.* and *Culex spp.* populations can have separate population dynamics, particularly with regards to environmental factors (Ezanno, et al., 2015).

As is to be expected, precipitation (Hasyim, et al., 2018; Thomson, Mason, Phindela, & Connor, 2005; Sena, Deressa, & Ali, 2015; Walldorf, et al., 2015), elevation (Tonye, Kouambeng, Wounang, & Vounatsou, 2018; Roberts & Matthews, 2016; Hasyim, et al., 2018; Daygena, Massebo, & Lindtjørn, 2017; Brooker, et al., 2004; Bødker, et al., 2003), and temperature (Thomson, Mason, Phindela, & Connor, 2005; Midekisa, Beyene, Mihretie, Bayabil, & Wimberly, 2015; Sena, Deressa, & Ali, 2015) all show associations with malaria.

2.2 Cluster-Randomized Control Malaria Trials

2.2.1. Effects of Malaria Vector Control Interventions on Malaria

Household-level usage of preventive measures is generally associated with reductions in malaria prevalence, although these associations may be confounded by several factors. While the use of ITNs and LLINs have been demonstrated to reduce the risk of malaria infection as discussed previously (Lengeler, 2009), ownership does not directly translate to use. Proper LLIN use requires that the net must be hung over a bed or other sleeping space and that individuals within the household must sleep under it each night. Studies of LLIN adherence during distribution programs and randomized controlled trials have shown that heat, social disruption, and the structural attributes of the sleeping structure may impact net use (Alaii, et al., 2003; Binka & Adongo, 1997; Raghavendra, et al., 2017). Studies which use LLIN ownership as a measure of LLIN use tend to find a non-significant association with malaria infection (Froeschl,

et al., 2018), and some studies have found no relationship between ITN use the previous night and malaria infection (Maziarz, et al., 2018; Bannister-Tyrrell, et al., 2018), although these results may be due to confounding by malaria endemicity as there is a strong correlation between malaria prevalence and net ownership and usage (Burgert, Bradley, Arnold, & Ekert, 2014; Atieli, et al., 2011).

Indoor Residual Spraying has shown to be an effective tool for the prevention of malaria, earning it a position alongside LLINs as one of the WHO-recommended malaria prevention methods (WHO, 2017). Studies on large-scale malaria programs or randomized controlled trials have consistently found malaria to be negatively associated with IRS (West, et al., 2013; Roberts & Matthews, 2016). However, studies that examine single-household IRS, rather than large scale initiatives often show no protective effect for the intervention (Maziarz, et al., 2018). This discrepancy may be due to the improper use of IRS during individual household spraying, information bias on the part of observational studies which fails to accurately capture the effect, or it may speak to a combined community effect on the part of larger scale interventions.

2.2.2. Evaluation of Vector Control Interventions

Cluster randomized controlled trials (cRCTs) are frequently used to evaluate the effectiveness of LLINs and IRS on clinical and entomological outcomes in many different malaria transmission settings. In these trials, individual villages are randomized as clusters, and all households within a given cluster are allocated to an intervention arm or control arm. This allows for easier distribution of the trial interventions, estimation of community effects, and protection from localized contamination between interventions. Cluster effects can then be

accounted for in the analytic models in order to control for similarities within clusters, although this increases the necessary sample size.

Most recent trials on IRS and LLIN interventions for malaria have opted for a cRCT design including trials in Ethiopia (Deressa, et al., 2014), Cambodia (Sluydts, et al., 2016), Kenya (Bousema, et al., 2016), and Tanzania (Mtove, et al., 2016). A previous cRCT has been run in Muleba, Tanzania on ITN and IRS interventions (West, et al., 2014). The WHO Vector Control Advisory Group guidelines for phase III vector control trials recommends either individual or cluster-randomized control trials, noting that cluster allocation helps prevent contamination between study arms and allows for the investigation of indirect or community level effects (VCAG, 2017).

2.2.3. Spillover Effect

Due to monetary and logistical concerns – and the political realities of delivering a large-scale intervention across a developing region – as well as a desire to have clusters between arms be as geographically similar as possible, trial organizers will often select villages within close proximity to one another for study. While this can cut down on costs and help reduce statistical differences between clusters, the community effects of bednet interventions discussed previously can become an issue when intervention and control arms are too close together. The effects of vector control interventions can result in localized reductions in malaria transmission that extend beyond range of the individuals using the intervention. This contamination or spillover effect between study arms can cause the intervention effects to become diluted. Silkey *et al*. suggest that it is possible to estimate the spillover effect of malaria bednet trials using the range and decay of effect as it approaches the cluster boundary (Silkey, et al., 2016). A 2003 study by

Hawley *et al*. in Western Kenya found that the community effect extended at least 300 meters, with some outcome measures showing a protective effect up to twice that (Hawley, et al., 2003). In order to prevent this, it has become common for cluster randomized trials to build a buffer zone surrounding cluster cores into their trial design with some buffers as large as 1500 meters (Bousema, et al., 2016). Delrieu *et al.* recommend a mixed buffer and gradient approach where sampling is conducted throughout the study area, but interventions are restricted to core areas allowing the decay effect to be assessed completely (Delrieu, et al., 2015). While this would increase the cost of cluster randomized trials, it would allow for much more accurate representations of intervention efficacy.

3. Methods

3.1 Study Data

Data used in this thesis come from the cluster-randomized trial *Evaluation of a Novel Long Lasting Insecticidal Net and Indoor Residual Spray Product, Separately and Together, Against Malaria Transmitted by Pyrethroid Resistant Mosquitoes*[4] commenced in Muleba, Tanzania in 2014 led by the Pan-African Malaria Vector Research Consortium (PAMVERC), a collaboration between the London School of Hygiene & Tropical Medicine (UK), Kilimanjaro Christian Medical University College (Tanzania) and the National Institute for Medical Research (Tanzania). The four-armed intervention study investigates the efficacy of a conventional Olyset LLIN, an Olyset Plus LLIN, a conventional Olyset LLIN in combination with a novel IRS compound, and an Olyset Plus LLIN in combination with a novel IRS compound. The conventional Olyset LLIN contains the current WHO recommended single pyrethroid (permethrin), while the Olyset Plus LLIN includes permethrin plus piperonyl butoxide (PBO), a synergist which slows the oxidative metabolic process in mosquitoes by stopping enzymes that would otherwise break down the permethrin (NPIC, 2017). The new IRS formulation is Actellic®300CS (micro-encapsulated pirimiphos-methyl), a non-permethrin based insecticide (WHO, 2013). Both the PBO and Actellic®300CS are used in an attempt to address pyrethroid resistance in the mosquito populations.

Forty-eight (48) clusters were identified in Muleba District, Tanzania, with 12 clusters assigned to each of the four study arms and an expected sample of 80 children (maximum six per

[4] ClinicalTrials.gov Identifier: NCT01697852

household) ages six months to 14 years per cluster. At the first post-intervention survey an additional 20 children were selected from households in the buffer areas for sensitivity analyses.

The obtained dataset includes an initial mapping survey that gathered geocoded household coordinates and data on basic household characteristics. Restricted randomization was done using measurements of malaria prevalence, mosquito and housing density, elevation, and net use and ownership before clusters were allocated to intervention arms. Data were then collected in cross-sectional surveys of random population subsets at baseline (September 2014) as well as four months (June 2015), nine months (November 2015), sixteen months (June 2016), and twenty-one months (November 2016) post-intervention. SES, IRS, and LLIN indicators were gathered at the household level in each post-intervention survey. The cross-sectional surveys include a follow-up clinical and parasitological component comprised of Rapid Diagnostic Testing (RDT) for malaria, temperature, and hemoglobin levels, for each child within a subset of randomly selected households. Only the initial mapping survey data and data from the baseline[5] and 4-month post-intervention[6] cross-sectional surveys were analyzed in this thesis.

All data collected in this trial are geocoded in World Geodetic System 1984 (WGS_1984) to allow for geospatial analyses. Initial mapping survey coordinates for each household were imported to ESRI ArcGIS 10.4 (ArcGIS, 2015) and projected into Africa Equidistant Conic projected coordinate system, preserving distances between points as accurately as possible for spatial analyses **(*Figure 3.1*)**. The spatial extent of the study covers 1433km^2 between 657,023m and 686,018m east of meridian 25 and 187,255m to 236,686m south of the equator and includes 29,311 surveyed households in the Muleba district of the Kagera region of Tanzania. Elevation

[5] The baseline cross-sectional were selected in order to capture the geographically-weighted odds of malaria infection in the absence of an intervention.
[6] The 4-months post-intervention was the only cross-sectional to gather data from non-core sampling households, which established the necessary sample size for semivariance analysis.

of study households ranges from 1075m to 1654m above sea level. A series of hills bisects the study region with low-lying households occupying the western areas and along the shore of Lake Victoria in the southeast **(*Figure 3.2*)**.

3.1.1. Baseline Study Population

The baseline household survey conducted in October 2014 recorded 2,880 households, 2,856 of which matched to households in the initial mapping survey, permitting geocoding of the data **(*Figure 3.3*)**. A total of 3,871 children completed a following clinical visit, of which 3,801 were able to be matched back to baseline household survey data for geocoding. Two-hundred fifty-nine (259) children were missing data on response or explanatory variables used in the principal component or logistic regression analyses, with 156 of those missing data on their RDT status. The final analyzed dataset comprised 3,542 children from 1,889 households.

3.1.2. Post-Intervention Study Population

The 4-month post-intervention household survey conducted July 2015 recorded 3,316 households, 3,282 of which matched to households in the initial mapping survey, permitting geocoding of the data **(*Figure 3.4*)**. A total of 8,019 children completed a following clinical visit, of which 7,635 were able to be matched back to post-intervention household survey data for geocoding. One thousand, one hundred and ninety-five (1,195) children were missing data on response or explanatory variables used in the principal component or logistic regression analyses, with 1,020 of those missing data on their RDT status. The final analyzed dataset comprised 6,440 children from 2,785 households.

3.2 Building Variables

3.2.1. Wealth Index

A principal component analysis was conducted using data collected during each household survey to build a measure of relative socioeconomic status for each household in the study population. Physical assets in the form of radios, mobile phones, bicycles, motorcycles, cars or trucks, canoes or other boats, and sewing machine, as well as household characteristics: number of rooms, whether the walls are plastered or bare, presence of a ceiling, electricity, farmland ownership, and livestock were all included in the initial PCA. These variables are all indicated as useful for determining socioeconomic status by the DHS (Vyas & Kumaranayake, 2006). Additional variables usually included in DHS surveys, sanitation facilities, water source, televisions, refrigerators, etc., were not included in the household surveys and as such could not be included in the PCA. Some variables such as roofing material and head of household level of education were purposely not included in order to minimize any collinearity between variables in the final risk factor model.

The PCA is calculated by creating a series of eigenvectors, or multidimensional perpendicular vectors that explain the greatest amount of variance possible for the given set of correlated variables in as many dimensions as there are variables by providing linear sums of the common factors for the variance of those variables (Rao, 1964). If you imagine a three-dimensional cloud, then a set of eigenvectors would be the three vectors, drawn perpendicular to each other in the x, y, and z dimensions, that best explain the variance of the cloud in each dimension by combining correlating variances between the variables or "principal components." In SAS v9.4 the calculated principal components are the least-squares solution to the linear model:

$$y_{ij} = x_{iq} b_{qj} + e_{ij}$$

(SAS, 2014)

where y_{ij} is the value of the ith observation of the standardized variable j (in this case durable assets), x_{iq} is the correlating value of the ith observation contributing to the principal component q, b_{qj} is the regression coefficient (or eigenvalue) of the common factor q predicting j, and e_{ij} is the residual error for the ith observation on the jth variable while ensuring that components have variances equal to their eigenvalues. The first principal component is predominantly considered to be an accurate representation of household economic status in the literature (Houweling, Kunst, & Mackenbach, 2003; Vyas & Kumaranayake, 2006).

3.2.2. Distance to Cluster Median Center

Distance to the cluster median center was used as a measurement of population density and continuous measure of rural/urban household status. The distribution of households in the study area tend to be spatially amorphous, making it difficult to identify a specific rural/urban divide. The cluster median center is a measure of central tendency which calculates the geographic center of the grouping of households in each trial cluster by minimizing the Euclidian distance to all households in the cluster (ESRI, 2018). Distance calculations to these points provide a reasonable approximation of the distance to the area of highest household density, or village center. Median center values were calculated based on the initial mapping survey so as to include all households in the study area, regardless of whether they were sampled in the cross-sectional surveys. The Median Center tool was used in ArcGIS 10.4 to calculate the center of each cluster and the Point Distance tool was used to determine the distance between each household and its respective center.

The formula for calculation of median center values is given as follows:

$$d_i^t = \sqrt{(X_i - X^t)^2 + (Y_i - Y^t)^2 + (Z_i - Z^t)^2}$$

(Burt & Barber, 1996)

where d_i^t is the Euclidian distance between locations i (the provided cluster datapoints) and t which is an iteratively generated median center candidate. X_i, Y_i, and Z_i are the x, y, and z coordinates for location i, X^t, Y^t, and Z^t are the x, y, and z coordinates for location t.

3.3 Geographically-Weighted Regression

3.3.1. Investigating Malaria Risk Factors

Malaria has been shown to have a spatial variability to its risk factors (Ndiath, et al., 2015; Homan, et al., 2016). That is to say, the impact of a risk factor can change based on the geographical location. This can be a change in the size of the effect, with some variables being more or less impactful in different areas, or a change in direction entirely where variables that are a risk factor in one part of the study area have a protective effect in others. *Ndiath et al* show this in their Geographically Weighted Regression analysis in Senegal, where household and village size both have local regression coefficients crossing 0 (Ndiath, et al., 2015), and *Manyangadze et al* show a similar relationship in Zimbabwe with temperature and altitude (Manyangadze, Chimbari, Macherera, & Mukaratirwa, 2017).[7]

Geographically Weighted Regression Analysis is primarily used to characterize and explain spatial patterns in data (Brunsdon, Fotheringham, & Charlton, 1996). In particular it can be used when there is an expectation that a predictor variable will have different effects across

[7] Fundamentally these changes in direction show the statistical anomalies in risk factors effects in an area that is wholly dominated by another factor (e.g. slight variability in age is unlikely to make a difference in odds of malaria infection when one lives in a marshland).

the study area. Studies have used Geographically Weighted Regression Analysis to establish geospatial trends in retailer distribution of anti-malarial medications (Rusk, et al., 2016) as well as geographic patterns in risk factors for diseases including Hepatitis C virus (Zhou, et al., 2016), HIV (Zhou, et al., 2015), AIDS (Alves, Nobre, & Waller, 2016), Hand-Foot-and-Mouth Disease (Zheng, et al., 2014), leptospirosis (Mayfield, et al., 2018), and malaria (Homan, et al., 2016; Ndiath, et al., 2015; Manyangadze, Chimbari, Macherera, & Mukaratirwa, 2017). Additional studies have shown Geographically Weighted Regression Analysis to perform better than the previously preferred method of Ordinary Least Squares (OLS) while allowing for spatial nonstationarity (i.e. variance in statistical relationships as a function of geographical location), assuming a large enough sample ($n\sim160$) (Devkota, Hatfield, & Chintala, 2014; Hasyim, et al., 2018).

The introduction of clusters into the regression model complicates the Geographically Weighted Regression analysis as the cluster weights matrix would explain some of the spatial trend that the analysis is attempting to define. Methods used to account for local variation in a Mixed Geographically Weighted Regression Analysis have been defined (Brunsdon, Aitkin, Fotheringham, & Charlton, 1999; Nakaya, Fotheringham, Charlton, & Brundson, 2009) including discussion for assigning terms to local or global (Brunsdon, Fotheringham, & Charlton, 1999; Mei, He, & Fang, 2004) and used to some success with dengue hemorrhagic fever (Astuti, Saputro, & Susanti, 2017), and tuberculosis (Octavianty, Jaya, & Toharudin, 2017). A similar semiparametric method was proposed by Huque *et al.* who discussed its application on ischemic heart disease (Huque, Bondell, Carroll, & Ryan, 2016). These methods attempt to control for the spatial variation within the model, rather than describing it by building the global fixed effects into the model. While this is useful for predictive modelling or extracting

effects for modelling residuals (as will be discussed later), it is unhelpful for characterizing spatial trends in the study area.

3.3.2. Analytic Approach: Geographically-Weighted Regression Analysis

The primary objective of Chapter 1 is to describe and characterize any spatial trends in the prevalence of malaria infection and identify determinants of malaria infection in children six months to 14 years of age at baseline, prior to intervention roll out.

Descriptive statistics were calculated for all variables of interest in the baseline clinical cohort (i.e. the sample of children tested for malaria by RDT in the baseline cross-sectional survey). Age, household elevation, distance of household from cluster center, presence of open eaves, type of roofing material, head of household schooling, wealth quintile, reported bednet use, and prior household IRS spraying were all expected to be significantly associated with malaria infection in children due to findings from a review of extant literature. Univariable associations with odds of malaria infection, defined as children that tested positive for *Plasmodium spp.* parasites via rapid diagnostic test, were determined through logistic regression using SAS 9.4 (SAS, 2012). Variables with a significant association were selected for a preliminary multivariable logistic regression model establishing the odds of malaria infection at baseline. All selected variables were checked for spatial autocorrelation through the calculation of a Moran's I statistic using SAS 9.4, which evaluates whether the pattern expressed is clustered, dispersed, or random.

A Kulldorff spatial scan statistic was calculated using SaTScan v9.4.4 (SaTScan, 2016) to determine the presence of high or low rate clusters for each of the predictor variables of interest. This creates a testing window centered on each household in the dataset and generates

Monte Carlo replications (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) to indicate whether the presence of a variable of interest is more or less likely in the defined area than randomly generated datasets. Dichotomous data was assessed using the Bernoulli model (Kulldorff, 1997), categorical data was assessed using the Ordinal model (Jung, Kulldorff, & Klassen, 2007), and continuous data was assessed using the Normal model (Kulldorff, Huang, & Konty, 2009). A Bonferroni correction (Bonferroni, 1936) of 1,889 was applied to the p-values of clusters detected by the Bernoulli model in order to control for the 1,889 household-centered windows tested in the SaTScan analysis. Clusters from the Ordinal and Normal models with a significance of $p<0.05$ were retained and described. Baseline distribution and significant clusters for each variable were mapped in ArcGIS 10.4.

GWR4 (GWR4, 2015) software was used to calculate a multivariable geographically-weighted logistic regression model. An adaptive bi-square distribution was fitted to the model data for each datapoint, weighting the measurements of study datapoints by their proximity to the point in question. The bi-square or biweight distribution is characterized by a gaussian-proximal center, but with an absolute cutoff rather than a continuous decline in the tails (Hoaglin, Mosteller, & Tukey, 1983). This absolute cutoff improves the resistance to outliers in the dataset and is recommended for regression data with a relatively high number of events (Nakaya, GWR4.09 User Manual, 2016) as it is able to make more localized estimates than a continuously smoothed regression. The cutoff bandwidth was determined by golden section search. This iterative mathematical process attempts to determine the maximum or minimum of a function (in this case the AICc of the k nearest neighbours) by calculating a set of three points from the input data: the maximum, minimum, and some point between them. Once the functions of these inputs have been determined, a fourth is selected and based on the function output either the top or

bottom spectrum of the input values are discarded and the process repeats based on the new range until the minimum value is found (Kiefer, 1953).

The geographically weighted logistic regression function used is given as follows:

$$y_i \sim Bernoulli[p_i]$$

$$logit(p_i) = \sum_k \beta_k(u_i, v_i) x_{k,i}$$

where p_i is the probability of the dependent variable for location i (in this case malaria infection), $x_{k,i}$ is the distance-weighted matrix of explanatory variable k given location i, β_k is the coefficient of explanatory variable k varying based on (u_i, v_i), the geographic coordinates for location i (Nakaya, 2016).

The adaptive bi-square bandwidth function (***Figure 3.5***) used is given as follows:

$$w_{ij} = \begin{cases} \left(1 - \frac{d_{ij}^2}{\theta_{i(k)}}\right)^2 & d_{ij} < \theta_{i(k)} \\ 0 & d_{ij} > \theta_{i(k)} \end{cases}$$

where w_{ij} is the weighted value of observation j for location i, d_{ij} is the fixed Euclidean distance between locations i and j, and $\theta_{i(k)}$ is the adaptive bandwidth size determined by the golden search distance of kth nearest neighbors (Nakaya, 2016).

A Normal model spatial scan statistic in SaTScan v9.4.4 was again used to determine clusters of high and low odds ratio concentration output by the geographically-weighted logistic regression described above. The results of the regression analysis and cluster detection were mapped in ArcGIS 10.4 to graphically display the spatial trends of each risk factor across the study area at baseline.

3.4 Semivariance Analysis

3.4.1. Community Effect Ranges

The impact of spillover effects on cluster randomized trials has been identified as an issue of concern by the WHO, which recommends several methods that studies can take to reduce spillover including buffer zones and "fried egg" cluster configurations, and cluster-randomized stepped-wedge designs (VCAG, 2017). While these recommendations suggest the need to contain any contamination effect, they do not offer suggestions on how large those areas should be to encompass the effects.

Several studies have attempted to examine theses ranges in various ways. A traditional approach to exploring the community effect of interventions is to measure the effects in close proximity to the cluster center and compare them to effects at greater distance (Bousema, et al., 2016). This proportion of effects gives a crude assessment of whether the effect greatly diminishes between the two distances, but does not allow for a further examination of the precise range of effect. A recent systematic review (Jarvis, Di Tanna, Lewis, Alexander, & Edmunds, 2017) identifies vector control studies that examine the contamination effect using comparisons across straight line distance (Hawley, et al., 2003; Binka, et al., 1998) and density (Lenhart, et al., 2008). While these examinations give more precise estimations of effect ranges, they still lack continuous measurement of range estimates. Silkey *et al.* utilize a susceptible-infected-susceptible model of disease transmission to simulate the potential for spillover effects in cluster-randomized trials, recommending that studies evaluate the "zones of imperfect coverage" to evaluate the range of effects more accurately (Silkey, et al., 2016). While this method does provide a continuous measure of simulated contamination effect range, it would require a drastic restructuring of the current method in which vector control trials are conducted. While it may be

found that the only way to model these ranges accurately is a major disruption in the methods of cluster-randomized vector control trials, the hope is to find a less intrusive way of establishing the presence of any contamination between clusters.

In order to obtain a continuous measure for the range of effect in vector control trials the second paper in this thesis constructs a fitted semivariogram model for each trial cluster using the residual odds of malaria from a geographically-weighted logistic regression model using likely confounding variables. Tobler's first law of geography states that correlation is a function of distance, with near samples tending to be more similar than distant ones (Tobler, 1970). A semivariogram is a model of this spatial autocorrelation in which the semivariance of measurements, or one half the mean squared difference in a set of measurements from two locations, are plotted against the distance separating those points (Cressie, 1993). This allows for a semivariogram curve to be fitted, originating at the origin and extending upwards as a function of distance.

There are three main components of the fitted semivariogram (***Figure 3.6***) that provide information on the spatial correlation of semivariance estimates as a function of distance. The Sill is the total scale of the semivariance in the study area; or more accurately the limit of the fitted semivariogram function as the model tapers off to a global value. The distance value at which the semivariogram model reaches[8] the Sill is called the Range, and the local effect of the data is the area to the left of this point. Finally, the Nugget is the value of any discontinuity in variance the fitted semivariogram model exhibits at the origin. This value is usually considered to be an estimation of measurement error in the dataset or evidence of microscale effects in the measurement of interest (Cressie, 1993).

[8] As the Sill is often approached asymptotically "reaches" has several definitions depending on the model chosen.

One additional parameter of interest can be calculated from the fitted model: the Partial Sill (sometimes referred to as the scale) is the total scale of the semivariance in the study area, this time discarding the portion of the variance present in the nugget discontinuity. This partial sill can be considered to be the value of the spatial variance in the dataset, as opposed to the nugget which is the value of the non-spatial variance, and is such the parameter of interest in this paper.

Semivariance models have been routinely used in spatial statistics to estimate continuous effect ranges in geoscience and geography for decades (Cressie, 1993; Banerjee, Carlin, & Gelfand, 2004). Outside these fields variogram modelling has been slowly expanding, with a wide range of applications. Studies have used semivariance modeling to determine the spatial autocorrelation of *Ixodes scapularis* ticks (Pardanani & Mather, 2004) and root knot nematodes (Oritz, Sullivan, Perry, & Vellidis, 2007), to assess the influence of environmental covariates on animal movement (Singer & Ovaskainen, 2017), to test for residual autocorrelation in an assessment of risk factors for Q fever-infected dairy cattle herds (Nusinovici, Frössling, Widgren, Beaudeau, & Lindberg, 2015), to explore spatial variation in the regression estimates of a geographically-weighted logistic regression model (Homan, et al., 2016), to model humidity from the conductivity of soil samples (Brom & Natonik, 2017), and to define the parameters for kriging of predicted Lyme disease transmission risk based on land use data (Nicholson & Mather, 1996), malaria incidence rates based on climate factors (Kleinschmidt, Sharp, Clarke, Curtis, & Fraser, 2001), and debris flow as predicted by precipitation (Yu, Xu, & Zhang, 2012). To the knowledge of this author, no studies have yet tried to estimate community effect ranges of intervention effects in cluster randomized trials by calculating the semivariance range.

3.4.2. Analytic Approach: Semivariance Analysis

The primary objective of Chapter 2 is to determine the presence of any community effects in the study area due to interventions, examine the range of these effects, and determine if they differ between intervention types.

Descriptive statistics were calculated for variables of interest in the 4-month post-intervention clinical cohort (i.e. the sample of children tested for malaria by RDT in the 4-month post-intervention cross-sectional survey). Age, sex[9], household elevation, distance of household from cluster center, presence of open eaves, type of roofing material, head of household schooling, wealth quintile were all expected to be significant predictors of malaria infection in children due to findings from Chapter 1 of this thesis. Univariable associations with malaria infection, defined as children that tested positive for malaria *Plasmodium spp.* parasites via rapid diagnostic test, were determined through logistic regression in SAS 9.4 (SAS, 2012). Variables with significant association were selected for a preliminary multivariable logistic regression model establishing the odds of malaria infection at baseline. All selected variables were checked for spatial autocorrelation through the calculation of a Moran's I statistic in SAS 9.4. GWR4 (GWR4, 2015) was used to calculate a multivariable geographically-weighted logistic regression (GWLR) model, using significant and spatially trending data from the global model,[10] with an adaptive bi-square bandwidth of the 439 nearest-neighbours using significant and spatially trending data from the global model. The adaptive kernel and golden section search were calculated as previously described. Residual values were extracted from the results of the GWLR for each datapoint and the dataset was subdivided into individual trial clusters. This division of

[9] Sex was not included in the baseline clinical questionnaire and was therefore only included as a covariate in Chapter 2.

[10] Distance from cluster center was omitted from this analysis as it is likely to capture most of the effect we are hoping to observe in the residual values.

the dataset into clusters is beneficial twofold: 1) it allows for the creation of quasi-stationarity in the cluster area as a sub-division of the greater study (Journel & Huijbregts, 1978), and 2) it allows for the assessment of cluster means individually, minimizing any spatial effect that could occur between clusters irrespective of cluster intervention.

For accurate estimation of variogram models, the study area must contain 50-100 data points, with 150-200 for good precision especially in the case of anisotropy (i.e. a state in which the underlying spatial process is directionally dependent) (Webster & Oliver, 1992). In addition to this it is recommended to have at least 30 (Journel & Huijbregts, 1978; Olea, 2006) or 50 (Chilès & Delfiner, 2012) pairs of data points in each lag group to ensure the accuracy of semivariance estimates with at least 3-5 lags (Lamorey & Jacobson, 1995) extending up to one half of the total bounded data distance in the study area (Journel & Huijbregts, 1978).

Traditional methods for fitting models to the semivariogram are highly subjective, referred to in some cases as being "as much art as science" (Banerjee, Carlin, & Gelfand, 2004). As this thesis paper is attempting to calculate range estimates across a mix of intervention groups it was determined that a unified approach should be taken in order to minimize the risk of bias in the results. As such, this paper utilizes a Weighted Least Squares (WLS) method to assess the goodness of fit for each variogram model (Jian, Olea, & Yu, 1996). Additionally, a robust estimator of the semivariance (Cressie & Hawkins, 1980) is used to provide stability in the replication of this method across each study dataset. Lag distances were calculated for each cluster to ensure a minimum of 50 data pairs per lag distance to produce a conservatively estimated semivariance (Chilès & Delfiner, 2012) and pairs whose paired distance exceeded one half the total bounded data distance in each cluster were excluded from this analysis (Journel & Huijbregts, 1978). A Moran's I was calculated in SAS 9.4 to assess the spatial autocorrelation for

the residual odds of malaria infection for each cluster dataset. The robust semivariance of this residual odds of malaria was then calculated in SAS 9.4 to examine residual spatial clustering in the data. A one-way analysis of variance (ANOVA) was conducted in SAS 9.4 to establish if there were any significant differences in the analysis characteristics for the clusters within each intervention allocation group.

A robust semivariance M-estimator Ψ for the average distance in class (h_k) is calculated in SAS 9.4 as follows:

$$\Psi(h_k) = \frac{1}{N(\theta_k, L)} \sum_{P_i P_j \in N(\theta_k, L)} \left[V(s_i) - V(s_j)\right]^{\frac{1}{2}}$$

(SAS, 2014)

where $P_i P_j$ is the pairing of points (in this case children who received diagnostic testing for malaria) i and j. These points are grouped into a population N, defined by an angle classification θ_k[11] and lag distance L. The summation of the square root of the difference in the values for the set of points $V(s)$ at locations i and j for population N is then averaged. Robust Semivariance estimates are then calculated in SAS 9.4 as follows:

$$\bar{\gamma}(h_k) = \frac{\Psi^4(h_k)}{2\left[0.457 + \frac{0.494}{N(\theta_k, L)}\right]}$$

(SAS, 2014)

where $\bar{\gamma}(h_k)$ is the robust semivariance estimate for the average distance in class h_k. A discussion of why the fourth-root M-estimator Ψ^4 provides a strong balance of stability and efficiency for most distributions can be found here (Cressie & Hawkins, 1980).

Robust semivariance estimates are then fit to the following forms in SAS 9.4:

[11] This analysis does not examine anisotropy in the study population and therefore the angle classification remains constant at 360 degrees. This ensures that no datapoints are omitted due to directionality.

Exponential $$\gamma_z(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ c_n + \sigma_0^2 \left[1 - \exp\left(-\frac{|h|}{a_0}\right)\right] & \text{if } 0 < |h| \end{cases}$$ (***Figure 3.7***)

Gaussian $$\gamma_z(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ c_n + \sigma_0^2 \left[1 - \exp\left(-\frac{|h|^2}{a_0^2}\right)\right] & \text{if } 0 < |h| \end{cases}$$ (***Figure 3.8***)

Spherical $$\gamma_z(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ c_n + \sigma_0^2 \left[\frac{3}{2}\frac{|h|}{a_0} - \frac{1}{2}\left(\frac{|h|}{a_0}\right)^3\right] & \text{if } 0 < |h| \leq a_0 \\ c_0 & \text{if } a_0 < |h| \end{cases}$$ (***Figure 3.9***)

Sine Hole Effect $$\gamma_z(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ c_n + \sigma_0^2 \left[1 - \frac{\sin\left(\frac{\pi|h|}{a_0}\right)}{\left(\frac{\pi|h|}{a_0}\right)}\right] & \text{if } 0 < |h| \end{cases}$$ (***Figure 3.10***)

(SAS, 2014)

where c_n represents the nugget effect, σ_0^2 is the partial sill (or scale), a_0 is the range estimate for each semivariance model form. Each semivariance model was compared in SAS 9.4 and selected by weighted least squared errors (WLS) to determine a model fit for each. The best fitting models by WLS were selected for each cluster unless identified as questionable by SAS 9.4. In these cases, the models were each examined in order of WLS until a suitable model was found. In cases where all 4 model structures were deemed questionable, a Power model was considered:

$$\gamma_z(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ c_n + \sigma_0^2 h^{a_0} & \text{if } 0 < |h|, 0 \leq a_0 < 2 \end{cases}$$ (***Figure 3.11***)

(SAS, 2014)

Significant range parameter estimates from the model of best-fit were used to calculate an "Effective Range" for each trial cluster equal to the a_0 parameter estimate for Spherical and Sine Hole Effect functions, $\sqrt{3}a_0$ for Gaussian functions, and $3a_0$ for Exponential functions.[12] These

[12] Gaussian and Exponential functions both approach the sill asymptotically and therefore do not have a "true" range estimate. As such, SAS 9.4 considers the Effective Range to be the point at which the function reaches 95% of the sill (SAS, 2014).

resulting Effective Ranges were grouped by intervention allocation for the four trial arms and compared to the 300 meter spillover buffer currently being used. Cases where WLS identifies a significant nugget parameter, but not a significant range estimate are considered to be a “Pure Nugget” effect (Oliver & Webster, 1986; Chilès & Delfiner, 2012) in which semivariance estimates remain reasonably constant as a function of distance and therefore have no local effect.

Figure 3.1 *Initial Mapping Survey projected in Africa Equidistant Conic projected coordinate system*

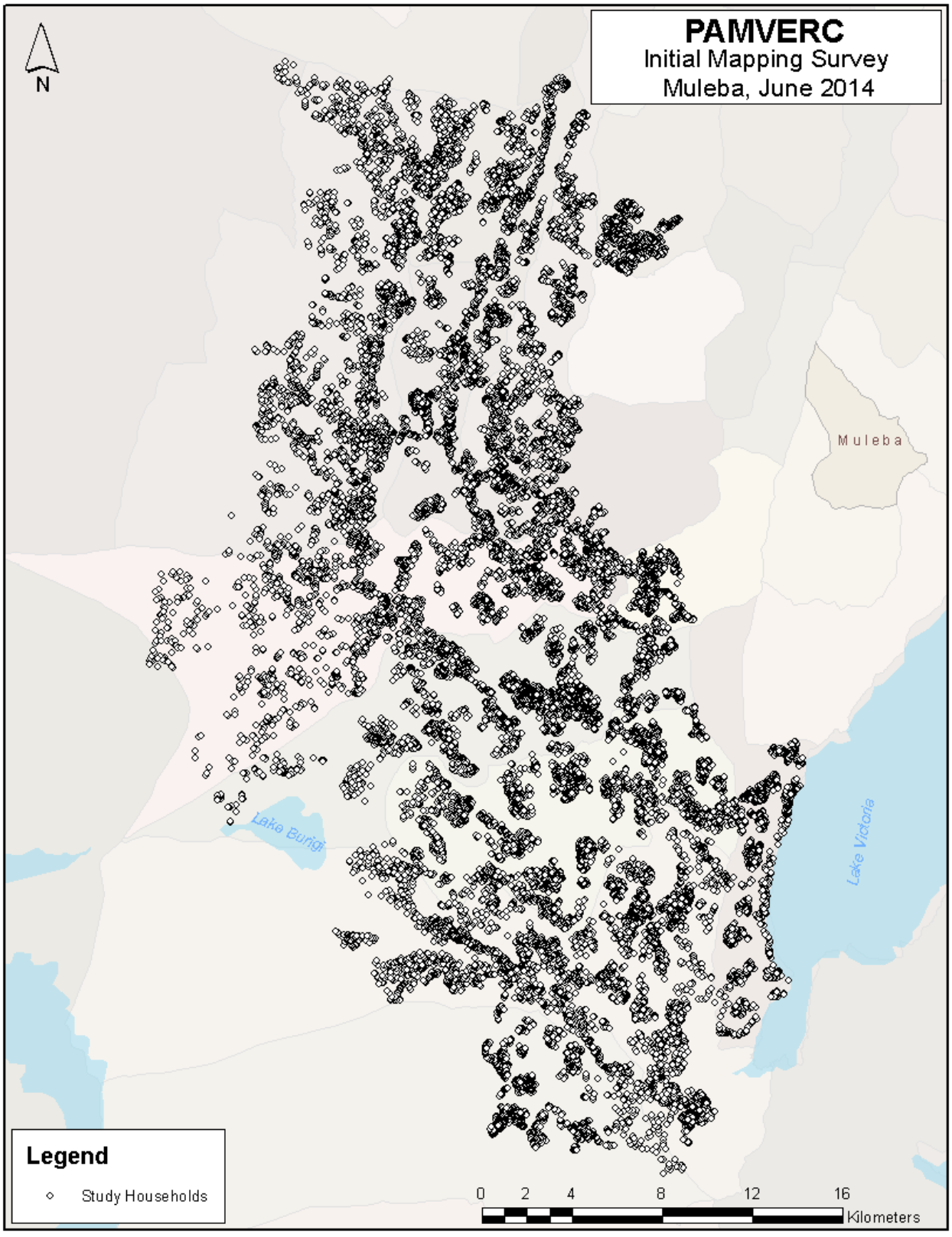


Figure 3.2 *Digital Elevation Model by Triangular Irregular Network*

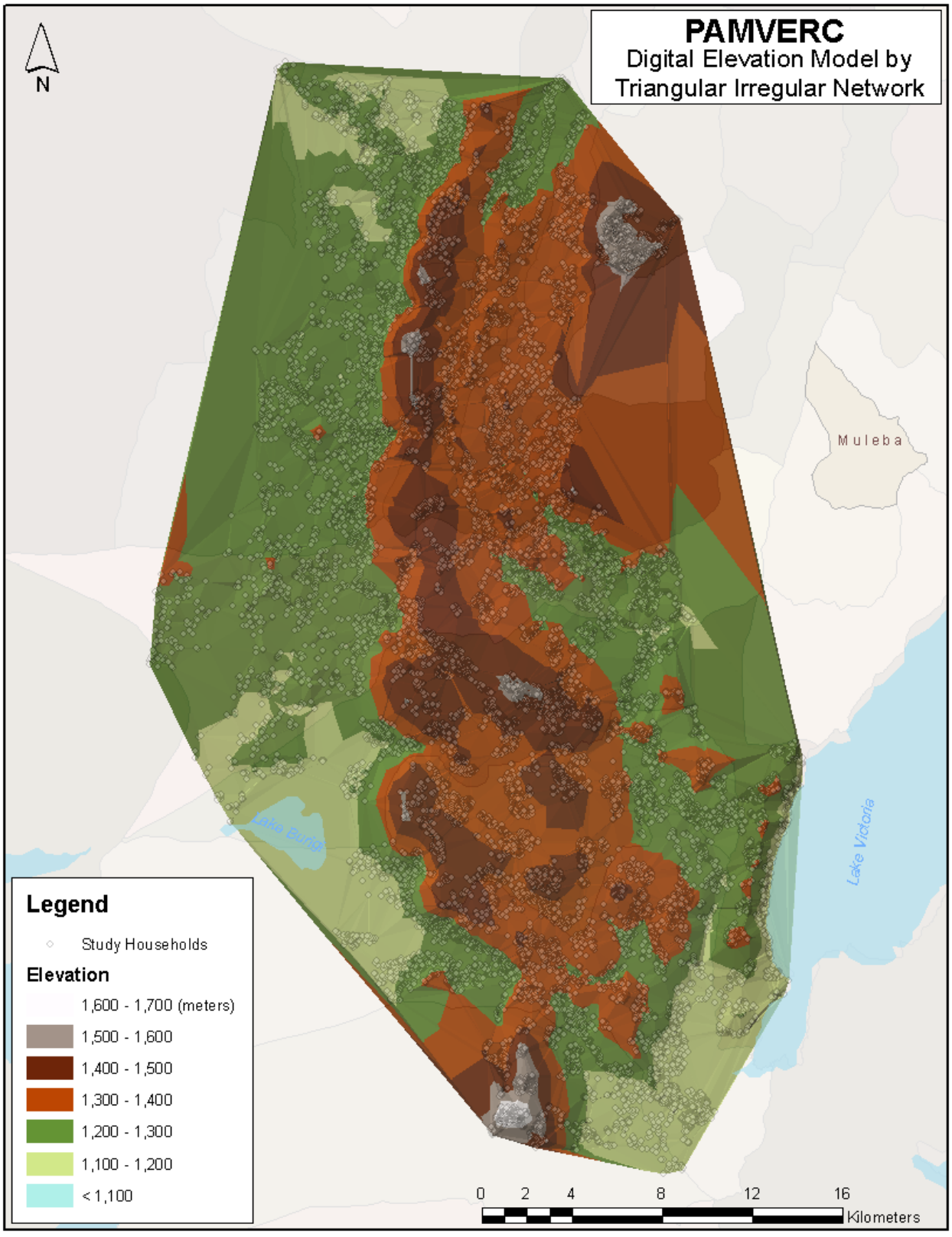


Figure 3.3 *Baseline Survey Muleba, October 2014*

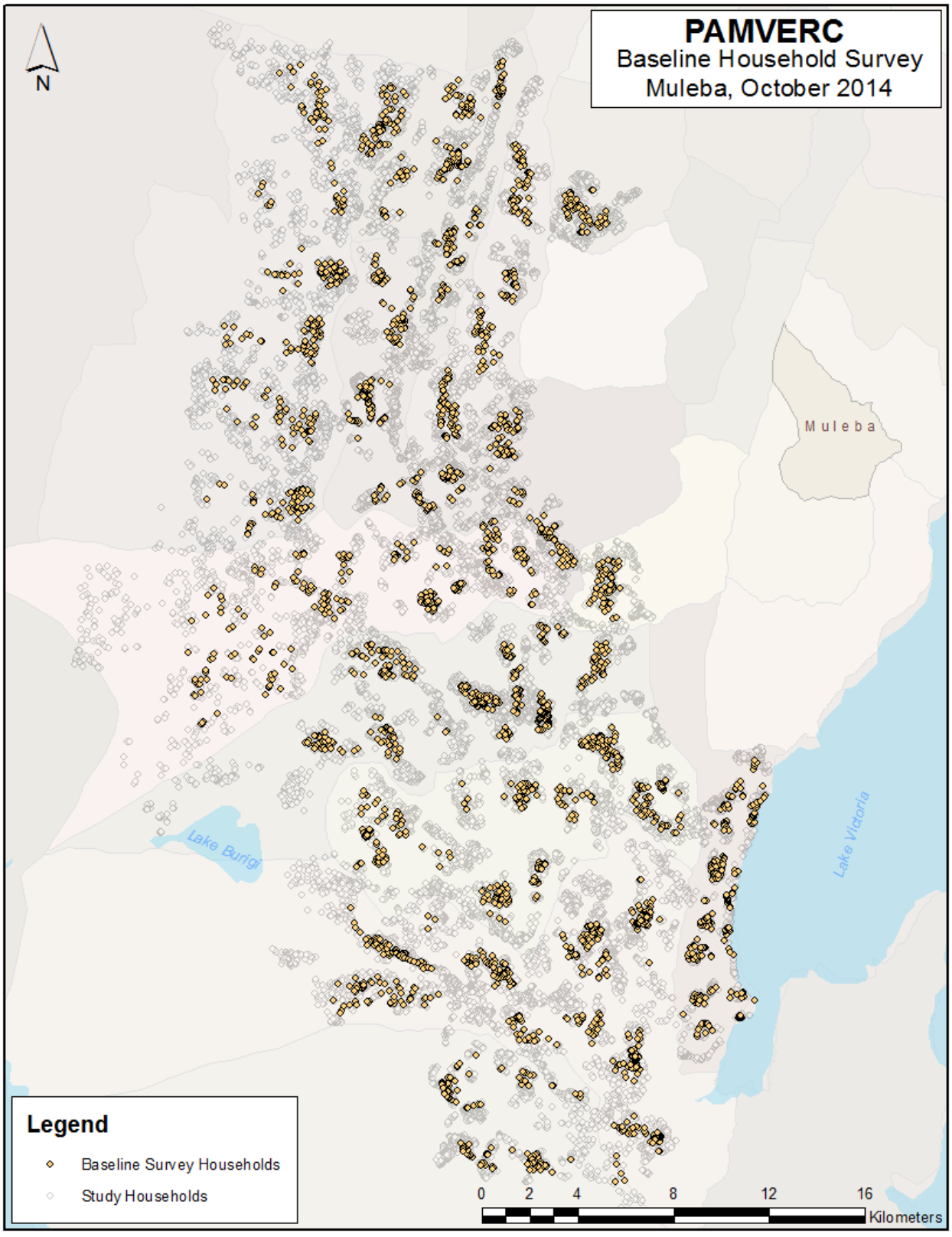


Figure 3.4 *Post-Intervention Survey Muleba, July 2015*

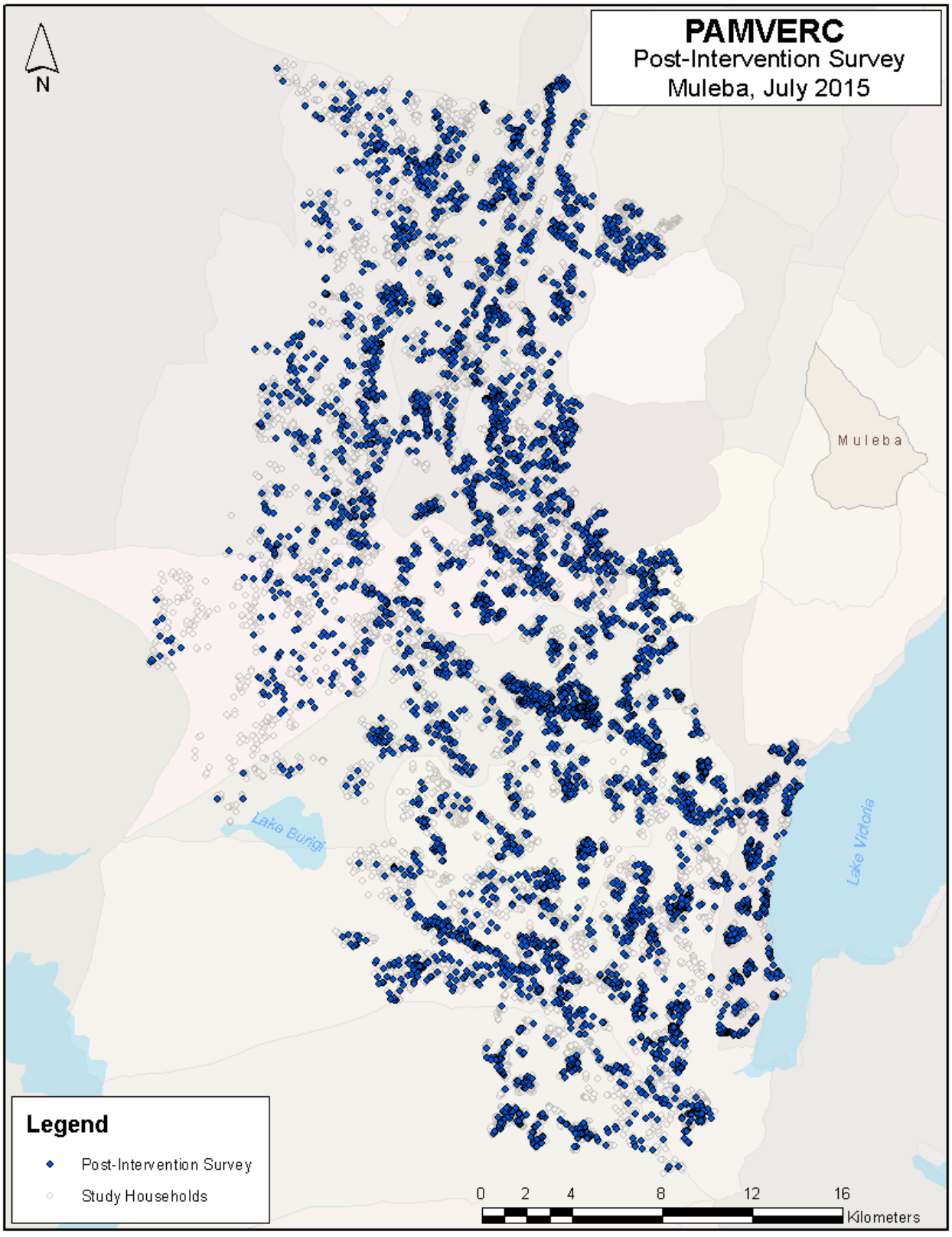


Figure 3.5 *Adaptive Bi-Weight Function for Geographically-Weighted Regression Kernel with Bandwidth of 1,000*

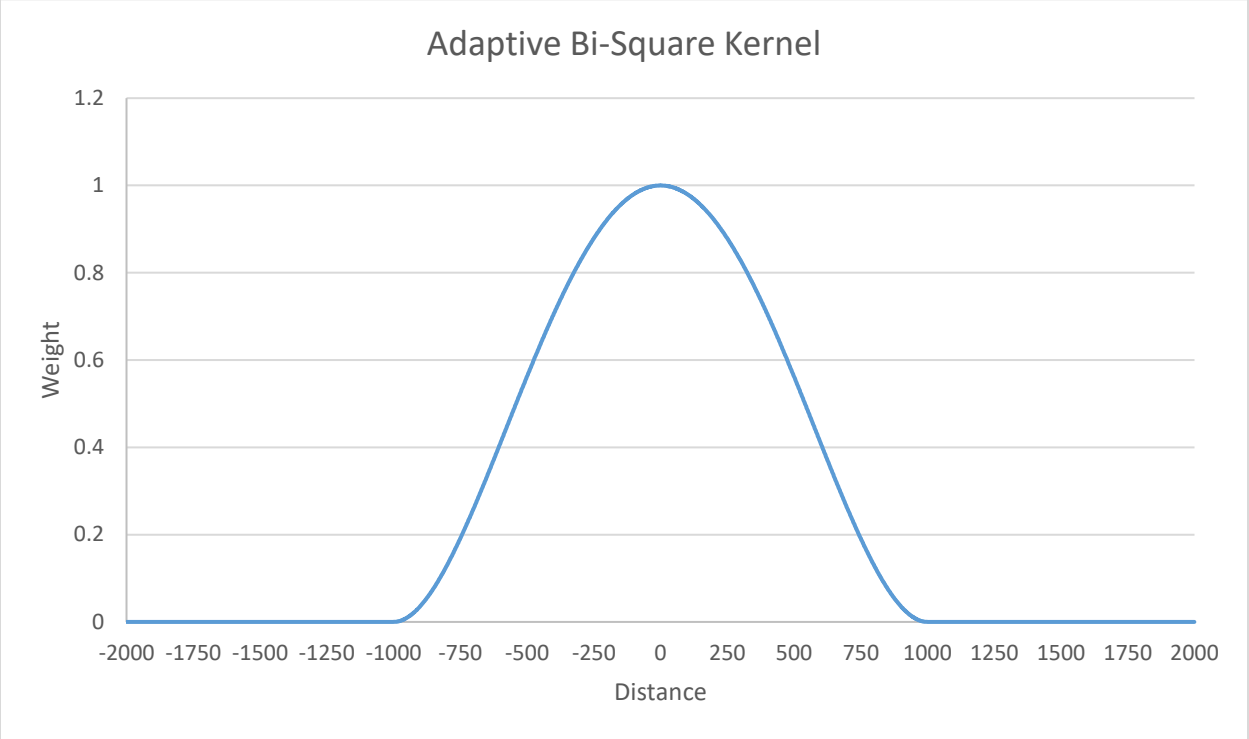


Figure 3.6 *Semivariogram Diagram*

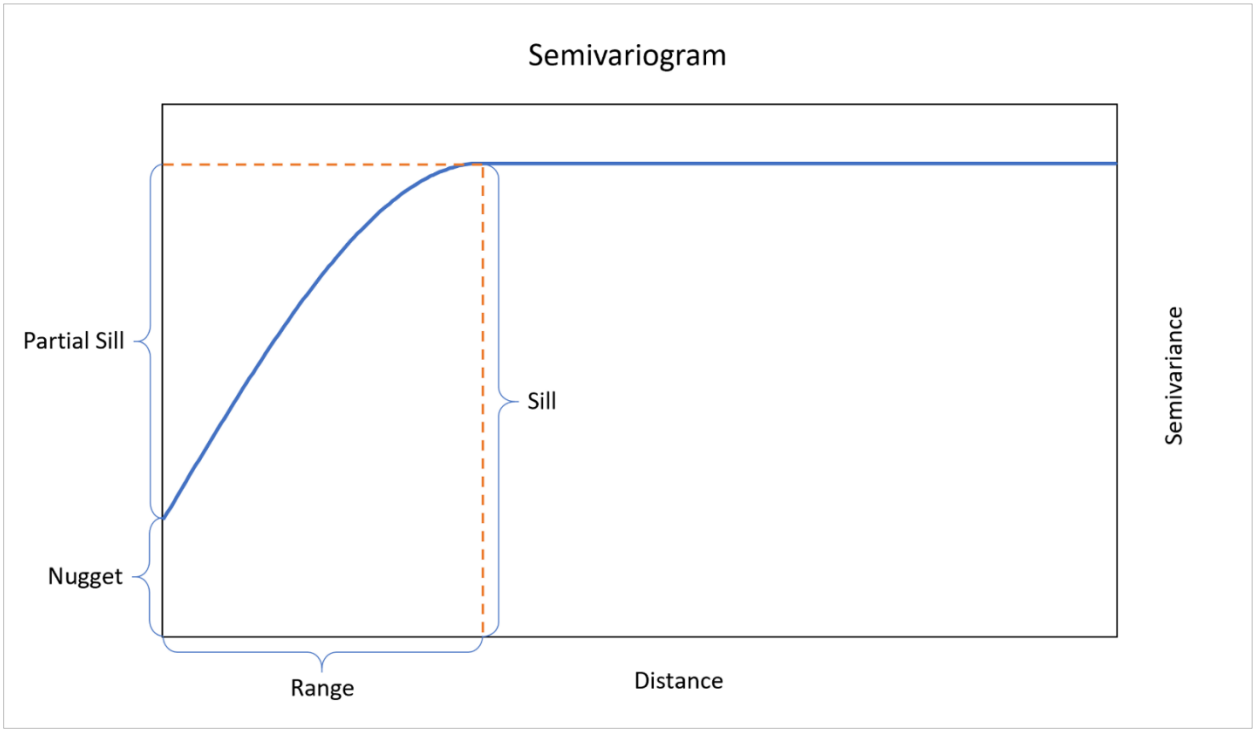


Figure 3.7 *Exponential Model for Semivariance Analysis with nugget (c_n) of 0, range (a_0) of 1, and sill (σ_0^2) of 4*

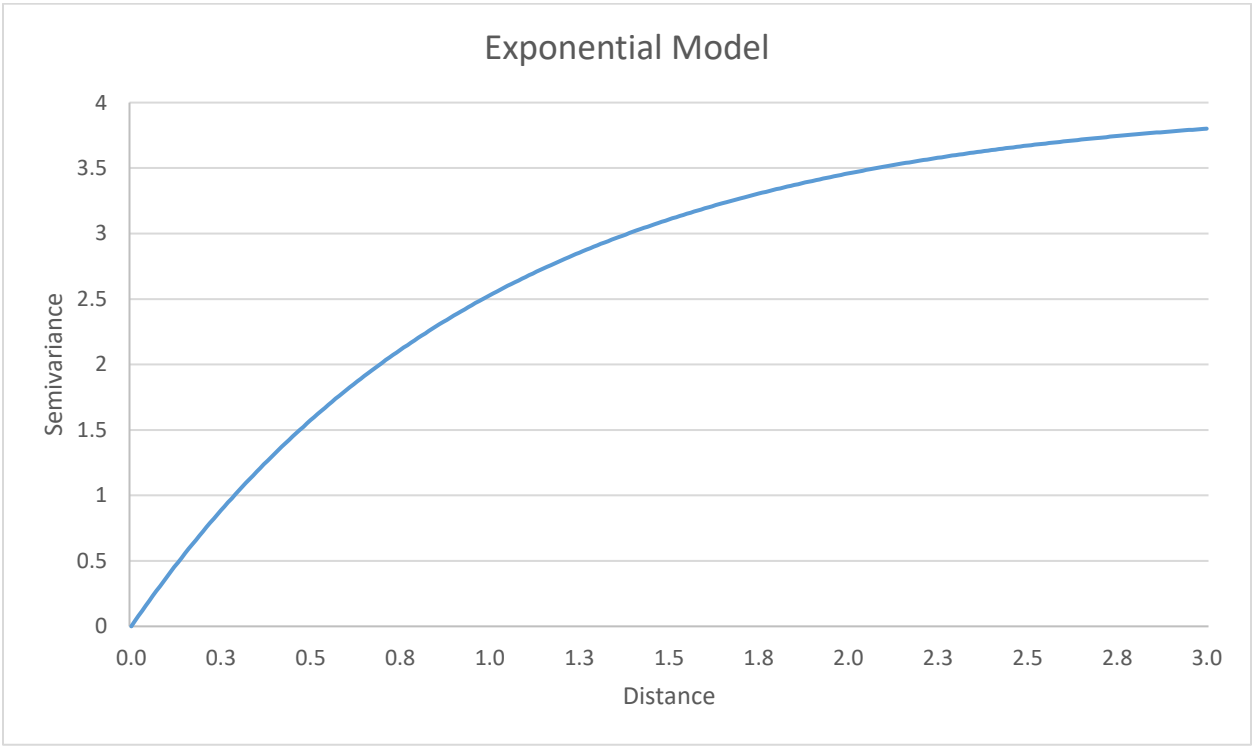


Figure 3.8 *Gaussian Model for Semivariance Analysis with nugget (c_n) of 0, range (a_0) of 1, and sill (σ_0^2) of 4*

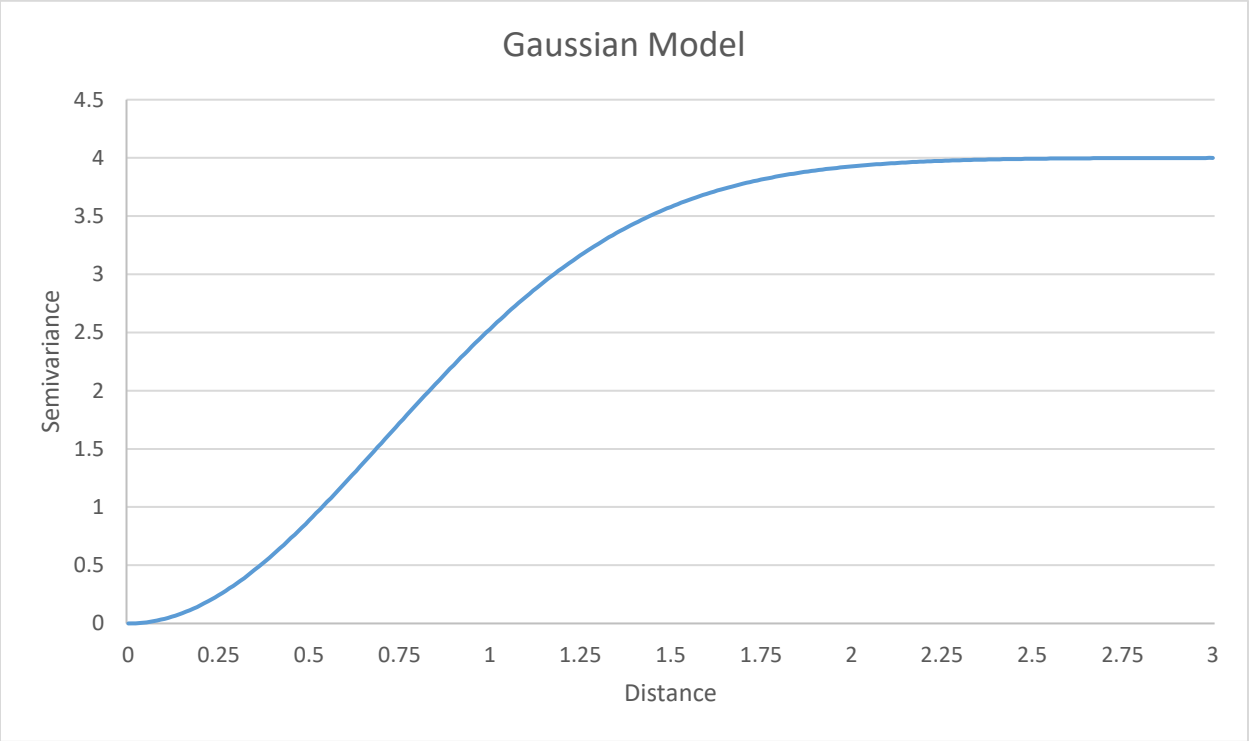


Figure 3.9 *Spherical Model for Semivariance Analysis with nugget (c_n) of 0, range (a_0) of 1, and sill (σ_0^2) of 4*

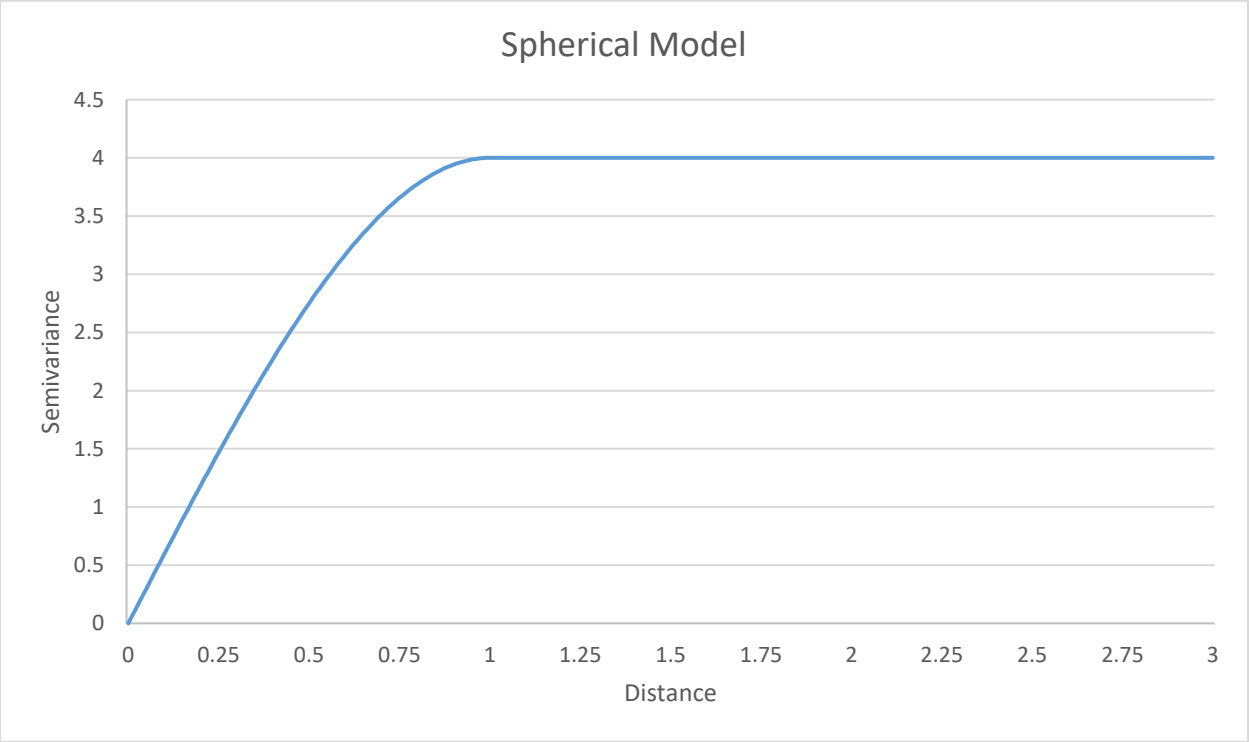


Figure 3.10 *Sine Hole Effect Model for Semivariance Analysis with nugget (c_n) of 0, range (a_0) of 1, and sill (σ_0^2) of 4*

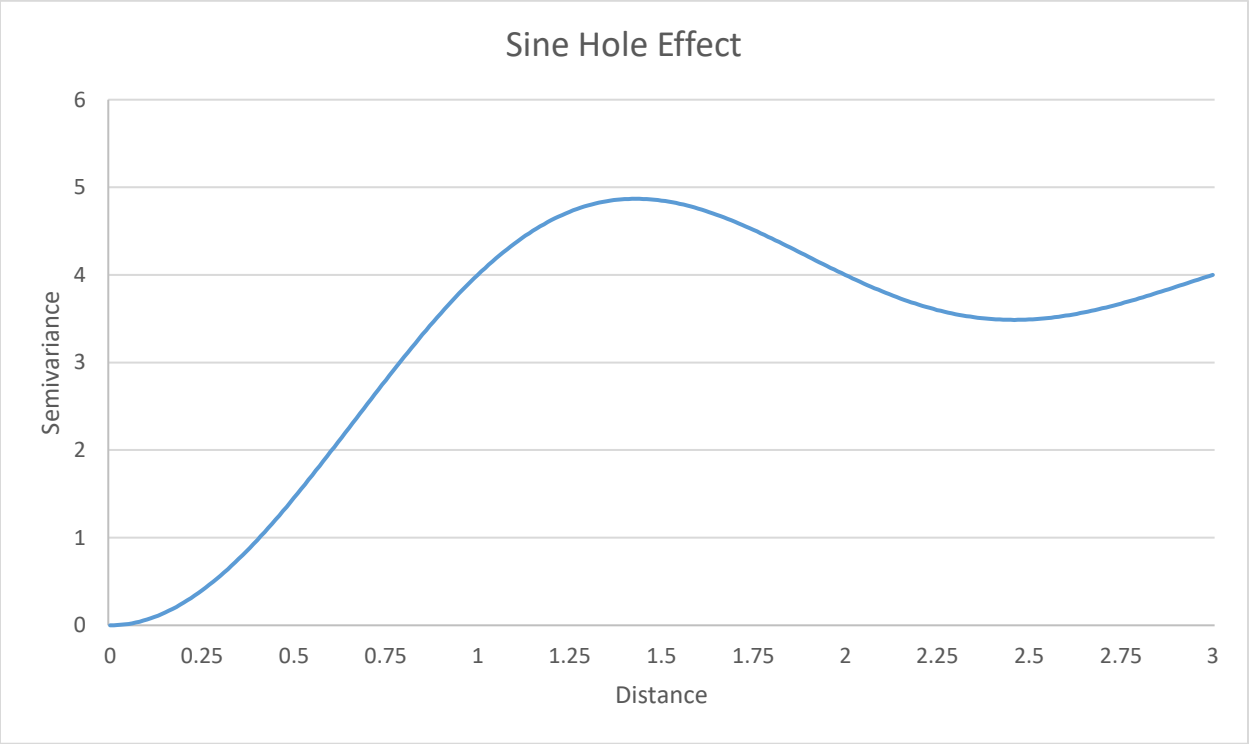
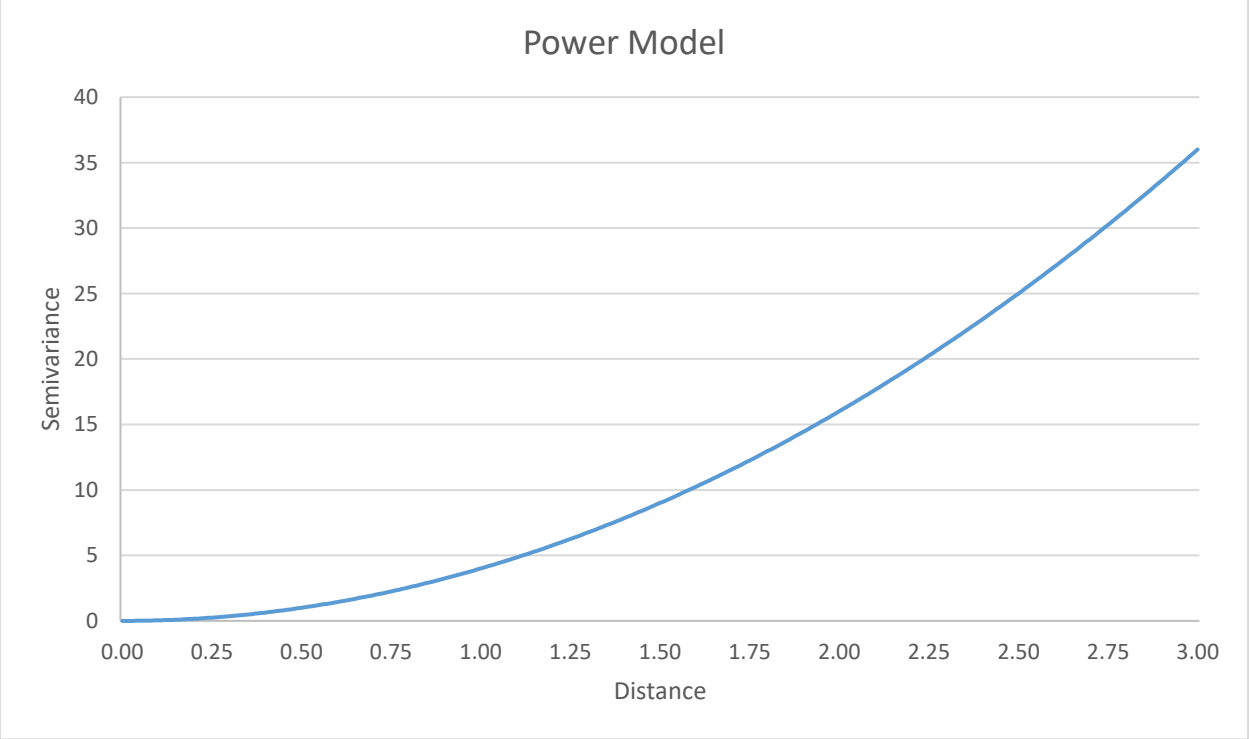


Figure 3.11 *Power Model for Semivariance Analysis with nugget (c_n) of 0, exponent (a_0) of 2, and slope (σ_0^2) of 4*



4. Chapter 1 – By All Odds

A Geographically-Weighted Logistic Regression Analysis of Risk Factors for Malaria in Muleba, Tanzania

Charles Russell Thickstun, Natacha Protopopoff, Ann Jolly, Alice Zwerling, Manisha Kulkarni

Formatted in preparation for submission to

The Malaria Journal

4.1 Abstract

Background While risk factors present a picture of what populations might be more likely to be infected with malaria, they are generally presented as being uniform across the entire study area. In the case of small-scale studies this is likely adequate, but previous studies have established that there is a spatial variability to malaria risk factors that could bias results across larger geographic areas.

Objectives This study aims to identify risk factors for malaria infection in children 6 months to 14 years and characterize the spatial variations of risk factor effects in Muleba, Tanzania.

Methods Univariable associations with odds of malaria infection were determined using standard logistic regression analysis of data from a household cross-sectional survey in Muleba, Tanzania. Variables with significant association were selected for a preliminary multivariable logistic regression model. All variables included in the model were checked for spatial autocorrelation. A multivariable geographically-weighted logistic regression (GWLR) model was subsequently calculated in GWR4 to describe spatial patterns in the odds of malaria

infection controlling for confounding variables. Cluster detection was used to identify spatial clusters in both variables and geographically-weighted adjusted odds of malaria infection.

Results Of 3,542 children tested by rapid diagnostic test, the prevalence of *Plasmodium falciparum* infection was 63.4%. All variables were spatially trending (Moran's I $p<0.0001$) and had odds that exhibited substantial spatial trends across the study area. All variables exhibited both risk-increasing and protective associations with malaria infection in different parts of the study area. One hundred meter change in elevation (Median 0.25 IQR: 0.12 – 0.46) and ITN/LLIN use (0.74 IQR: 0.48 – 1.25) demonstrated the greatest spatial variability as covariates in the GWLR model. One hundred meter change in distance from cluster center (0.99 IQR: 0.95 – 1.05), the household having open eaves (0.65 IQR: 0.39 – 1.00), natural roofing materials (0.90 IQR: 0.53 – 1.54), head of household education (0.87 IQR: 0.72 – 1.05), and wealth quintile all demonstrated large ranges in odds across the study area. One year change in age demonstrated the least change in odds (1.08 IQR: 1.03 – 1.15).

Conclusions Age of child and several household factors, such as elevation, type of eaves, roofing material, head of household education, wealth quintile, and household bednet use were all found to be associated with malaria in both standard and geographically-weighted logistic regression models. Both models produced similar OR for the covariates, but the GWLR model also detected spatial variation in the effect of each covariate. GWLR results could be used to inform spatially-targeted interventions for malaria to complement existing intervention strategies.

Keywords Malaria, Risk Factors, Geographically-Weighted Regression, Cluster Detection, Spatial Analysis

4.2 Background

Malaria is an acute febrile disease caused by *Plasmodium species* parasites that are transmitted by the bite of an *Anopheles species* mosquito (APHA, 2015). In 2017, 200 million cases and 403,000 deaths from malaria occurred in the World Health Organization (WHO) African Region (WHO, 2018) where *P. falciparum* predominates (Gething, et al., 2011). Sub-Saharan Africa accounts for 99.7% of all malaria cases (WHO, 2018), with *Anopheles gambiae* and *An. funestus* as primary vectors (Sinka, et al., 2012; Wiebe, et al., 2017). This is set against a backdrop of increasing insecticide resistance where the standard single-insecticide pyrethroid long-lasting insecticidal net (LLIN) fails to provide the coverage it once did, as the proportion of countries with endemic malaria transmission that both monitored and reported resistance to pyrethroid insecticides reached 81% in 2016 (WHO, 2017).

Risk factors associated with malaria infection or clinical cases have been explored extensively and act as a combination of factors influencing the *Plasmodium* parasite, the *Anopheles* vector, and the human host they come in contact with (Protopopoff, et al., 2009). As this is a complicated ecosystem of factors relating to the disease it can be difficult to single out exact effects without extensive testing of both vectors and hosts on multiple scales, which would prove prohibitively expensive. In order to establish meaningful associations with variables that can either be targeted for intervention or changed to provide protection from malaria, most studies have developed proximal variables or "risk factors" that are associated with malaria and can be assessed without enormous cost. In recent years major malaria studies have found associations with major demographic variables like age, socioeconomic status, housing quality, population density, and head of household education (Walldorf, et al., 2015; Roberts & Matthews, 2016; Bannister-Tyrrell, et al., 2018; Tonye, Kouambeng, Wounang, & Vounatsou,

2018; Were, et al., 2018; Homan, et al., 2016) and environmental factors like precipitation, elevation, and temperature (Hasyim, et al., 2018; Thomson, Mason, Phindela, & Connor, 2005; Sena, Deressa, & Ali, 2015; Walldorf, et al., 2015; Daygena, Massebo, & Lindtjørn, 2017; Midekisa, Beyene, Mihretie, Bayabil, & Wimberly, 2015; Sena, Deressa, & Ali, 2015). A previous study in Muleba, Tanzania found associations between *P. falciparum* infection and age, household socio-economic status, and malaria transmission intensity (West, et al., 2013).

While these risk factors present a picture of what populations might be more susceptible to malaria, they present a uniform result for the entire study area. In the case of small scale studies this is likely adequate, but previous studies have established that there is a spatial variability to malaria risk factors (Ndiath, et al., 2015; Homan, et al., 2016), which demonstrates that the impact or even direction of association between a variable and malaria infection can change based on the geographical location. This can be a change in the size of the effect, with some variables being more or less impactful in different areas, or a change in direction entirely where variables that are a risk factor in one part of the study area have a protective effect in others. Such a variation in risk could allow for interventions to be targeted specifically to regions where they would be most valued. This study characterizes the spatial variations of risk factor effects on malaria infection in Muleba, Tanzania using Geographically-Weighted Logistic Regression (GWLR) analysis to describe spatial patterns in the odds of malaria infection controlling for confounding variables.

4.3 Methods

4.3.1. Data Source

This study uses data from a baseline cross-sectional survey conducted in 2014 as part of a cluster-randomized trial in Muleba, Tanzania led by the Pan-African Malaria Vector Research Consortium (PAMVERC), a collaboration between the London School of Hygiene & Tropical Medicine (UK), Kilimanjaro Christian Medical University College (Tanzania) and the National Institute for Medical Research (Tanzania). The trial was comprised of 48 clusters, and the survey targeted 80 children (maximum six per household) ages six months to 14 years per cluster. Data were collected from a random sample of households using interviewer-guided questionnaires with questions on demographic, socioeconomic status (SES), Indoor Residual Spraying (IRS), and Long-Lasting Insecticidal Net (LLIN) indicators. The data includes a follow-up clinical and parasitological component which includes Rapid Diagnostic Testing (RDT) results for malaria, temperature, and hemoglobin levels, for randomly selected children within the selected households.

Data collected for the baseline household survey included only households in the core sampling areas designated for each cluster in the study – established as a minimum of 600m from another core sampling area (Protopopoff, et al., 2018).

4.3.2. Study Area and Population

Muleba district, Tanzania is situated along the western shore of Lake Victoria south of Bukoba. With a population of 540,310 Muleba is the largest district in the Kagera Region, and mostly rural as Muleba Town boasts a population of only 15,747 (National Bureau of Statistics, 2016). Data collected for the study were gathered from the western parts of the district,

excluding Muleba Town and its immediate vicinity. A previous study in the Muleba region found that *P. falciparum* infection in children aged 6 months to 14 years was 22.8% and varied dramatically across the study area, with highest levels in the southwest (West, et al., 2013) where the current study is located.

All data collected in this trial is geocoded in World Geodetic System 1984 (WGS_1984) to allow for geospatial analyses. Initial mapping survey coordinates for each household were imported to ESRI ArcGIS 10.4 (ArcGIS, 2015) and projected into Africa Equidistant Conic projected coordinate system, preserving distances between points as accurately as possible for spatial analyses. The spatial extent of the study covers 1433km^2 between 657,023m and 686,018m east of meridian 25 and 187,255m to 236,686m south of the equator and includes 29,311 surveyed households in the Muleba district of the Kagera region of Tanzania. Elevation of study households ranges from 1075m to 1654m above sea level. A series of hills bisects the study region with low-lying households occupying the western areas and along the shore of Lake Victoria in the southeast. The baseline survey used in this study was conducted in October 2014 and the final analyzed dataset comprised of 3,542 children from 1,889 households.

4.3.3. Statistical Analyses

Descriptive statistics were developed for all variables of interest. Child age, as well as household elevation, distance from cluster center, eave type, roofing material, wealth quintile, reported household bednet use the previous night, reported IRS spraying in the past year, and head of household level of education were all expected to be significantly associated with malaria infection due to findings from a review of extant literature. Univariable associations with malaria infection, defined as children that tested positive for malaria *Plasmodium spp.* parasites

via rapid diagnostic test, were determined through logistic regression in SAS 9.4 (SAS, 2012). Variables with significant association ($p<0.05$) were selected for a preliminary multivariable logistic regression model establishing the odds of malaria infection. All variables included were checked for spatial autocorrelation through the calculation of a Moran's I statistic in SAS 9.4.

A Kulldorff spatial scan statistic was calculated using SaTScan v9.4.4 (SaTScan, 2016) to detect high or low rate clusters for each of the predictor variables of interest. Dichotomous data was assessed using the Bernoulli model (Kulldorff, 1997), categorical data was assessed using the Ordinal model (Jung, Kulldorff, & Klassen, 2007), and continuous data was assessed using the Normal model (Kulldorff, Huang, & Konty, 2009). A Bonferroni correction of 1,889 was applied to the p-values of clusters detected by the Bernoulli model (Bonferroni, 1936). Clusters from the Ordinal and Normal models with a significance of $p<0.05$ were retained and described. Baseline distribution and significant clusters for each variable were mapped in ArcGIS 10.4.

GWR4 (GWR4, 2015) was used to calculate a multivariable geographically-weighted logistic regression model. An adaptive bi-square distribution was fitted to the model data for each datapoint, weighting the measurements of study datapoints by their proximity to the point in question. A Normal model spatial scan statistic in SaTScan v9.4.4 was again used to determine clusters of high and low odds ratio concentration output by the geographically-weighted logistic regression described above. The results of the regression analysis and cluster detection were mapped in ArcGIS 10.4 to graphically display the spatial trends of each risk factor across the study area at baseline.

4.4 Results

4.4.1. Characteristics of Study Population

Table 4.1 describes the characteristics of the study population. A total of 2,246 children tested positive for malaria parasites (*P. falciparum* or other) for an overall prevalence of 63.4%. Children who tested positive for malaria parasites were older on average (mean age 6.75 years, 95%CI:6.59 – 6.91) than those who tested negative (6.07 years, 5.85 – 6.28), lived at lower elevations (1,297m, 1,294 – 1,301 vs 1,345 meters, 1,340 – 1,351), and further away from cluster centers (988m, 960 – 1,015 vs 921m, 887 – 956). Malaria infection showed an association with households with open eaves (67.91% vs 56.32%, $p<0.0001$), natural roofing materials (74.70% vs 60.83%, $p<0.0001$), low head of household education level (70.25% vs 48.65%, $p<0.0001$), and low household wealth (71.15% vs 55.72%, $p<0.0001$). Malaria infection also showed an association with ITN or LLIN use (65.25% vs 60.67% $p<0.0056$), but not IRS (65.14% vs 63.08%, $p<0.3719$).

4.4.2. Spatial Characteristics of Baseline Study Population

Table 4.2 details the spatial trend of malaria infection and variables of interest. All variables of interest had a significant spatial trend in the study population (Moran's I $p<0.0001$). Significant clusters of high risk of malaria infection (Observed/Expected = 1.52 – 1.30) were concentrated in the central regions of the study area, with five low risk clusters in the southern regions (O/E = 0.54, 0.38, 0.35, 0.33, & 0.18) and two low risk clusters in the northern areas (O/E = 0.54 & 0.53) **(*Figure 4.1*)**. One high age cluster (O/E = 1.22) was identified in the northern region, and one low age (O/E = 0.79) in the southwest **(*Figure 4.2*)**. Three high elevation clusters (O/E = 1.20 – 1.07) draw a line down the middle of the center of the study

area, with low elevation clusters (O/E = 0.97 – 0.90) identified in the northwest and southeast along the shore of Lake Victoria **(*Figure 4.3*)**. Five clusters of rural households (O/E = 4.05 – 0.25) are fairly distributed across the western edge of the study area, with two in the north and center-east regions. Three clusters of urban households (O/E = 0.68 – 0.25) were identified, with two just west of Lake Victoria in the south, and one in the center-north area **(*Figure 4.4*)**.

Two clusters of households with open eaves (O/E = 2.57 & 1.59) occur in the southeastern regions northwest of Lake Victoria. Two clusters of households with closed eaves were identified in the far western region of the study area (O/E = 0.19) and the southern central region (O/E = 0.00) **(*Figure 4.5*)**. There are two clusters of roofs made of grass, leaves, or some combination of natural materials and metal (O/E = 4.56, 2.86) one in the westernmost region and one very small cluster in the south near the tip of Ruiga Bay, Lake Victoria. One cluster of roofs made with metal sheets sits in the northeastern most area (O/E = 1.12) **(*Figure 4.6*)**. Six high education clusters are spread across the eastern regions of the study area (O/E = 25.00 – 1.37), whereas two of the four low education clusters are found in the western center (O/E = 1.67) and center-north (O/E = 2.26). Smaller low education clusters can be found in the center-east (O/E = 3.70) and south (O/E = 3.69) **(*Figure 4.7*)**. Three of the six high wealth clusters (O/E = 4.65, 2.46, & 1.62) are located in the northern regions while the other half lay in the south (O/E = 4.65, 1.99, & 1.66). Unexpectedly low wealth clusters are found in the southeast along the shore of Lake Victoria (O/E = 1.79) and in the westernmost region north of Lake Burigi (O/E = 1.56) **(*Figure 4.8*)**. The three high bednet use clusters (O/E = 2.37 – 1.50) are concentrated in the southernmost region, and the two low bednet use clusters (O/E = 0.55 – 0.54) are in the northern central region **(*Figure 4.9Figure 4.10*)**. One very large high IRS cluster (O/E = 1.07) spans the

eastern central region of the study area, and two low IRS clusters (O/E = 0.67 – 0.56) are centered in the westernmost area just north of Lake Burigi **(*Figure 4.10*)**.

4.4.3. Global Logistic Regression

Standard univariable logistic regression models were constructed using SAS v9.4 to determine the odds of malaria infection as predicted by each variable (***Table 4.3***). All variables that showed significant associations by chi-square or t-test also displayed significant associations with logistic regression. Increased age was associated with a 5% increased odds of malaria infection (OR = 1.05, 95% CI: 1.03 – 1.06) per year. Houses further from cluster centers showed two percent increased odds of infection (OR = 1.02, 1.01 – 1.03) per hundred meters distance increase from cluster centers. Roofing materials which included grass or leaves were associated with almost twice the odds of malaria infection (OR = 1.90, 1.57 – 2.30) than those of all-metal construction. Decreased head of household schooling shows an increased odds of infection: no head of household education shows 2.5 times higher odds of infection compared to those with secondary/technical or higher education (OR = 2.50, 1.68 – 3.71), and 1.68 time higher odds of infection compared with primary education (OR = 1.68, 1.15 – 2.46). Household wealth shows a similar trend with the poorest houses having almost twice the odds of infection as compared to the least poor (ORs = 1.96, 1.57 – 2.45). Elevation showed a 43% decreased odds of malaria infection with every 100-meter increase in altitude (OR = 0.57, 0.53 – 0.62). LLIN/ITN use was also protective of malaria infection compared to children who reported not using a LLIN/ITN the previous night (OR = 0.82, 0.72 – 0.94). Application of IRS in the past year (2014) was not found to have significantly different odds of infection (0.92, 0.75 – 1.11; p=0.372).

The multivariable model included all variables that were significant in the univariable models age, elevation, etc. (***Table 4.3***). In this adjusted model, the odds of infection by age increased slightly (OR = 1.06, 1.04 – 1.08). All other variables, including elevation (OR = 0.56, 0.52 – 0.61), households with open eaves (OR = 1.45, 1.24 – 1.70), head of household education for both no education and primary education (ORs = 1.78, 1.16 – 2.71; 1.41, 0.95. – 2.11), all levels of wealth quintile (ORs = 1.30, 1.01 – 1.68; 1.45, 1.14 – 1.83; 1.06, 0.85 – 1.32; 1.07, 0.86 – 1.33), and ITN/LLIN use (OR = 0.78, 0.67 – 0.90) showed moderate decreases in odds of malaria infection in the adjusted model, but these associations were still significant. Distance from cluster centers (1.01, 1.00 – 1.02; p=0.1670) and roofing material (1.35, 1.07 – 1.72; p=0.0603) were no longer significantly associated with odds of malaria infection after controlling for other variables.

4.4.4. Geographically-Weighted Logistic Regression

The adjusted standard logistic regression model above was used to calculate the geographically-weighted adjusted odds of malaria for each child in the study. An adaptive bi-square bandwidth of 227 persons resulted in the lowest GWLR model AICc. The geographically-weighted regression model displayed a better fit than the global model (AICc 4,296 versus 3,133), indicating that the GWLR model was more efficient. Significant spatial clusters of high or low odds ratio values were detected for all variables. ***Table 4.4*** shows the geographically-weighted adjusted odds of malaria infection for each variable, including minimum, median and maximum OR values as well as interquartile OR values and overall range. Characteristics of the significant clusters of high and low odds ratio values for each variable and associated Moran's I

autocorrelation results are listed in ***Table 4.5***. The geographically-weighted adjusted odds for all variables had a significant spatial trend (Moran's I $p<0.0001$).

Age displayed an odds ratio range of 0.84 – 1.34 across the study region and maintained an increasing odds of malaria infection with increased age throughout most of the study area **(*Figure 4.11*)**. Three significant clusters of high OR for age (O/E = 1.20 – 1.10) were evident in the center-east, south, and southwest. Two clusters where age showed a statistically significant protective effect (O/E = 0.92 – 0.87) were apparent in the center and western regions. Household elevation had an odds ratio range of 0.00 – 29.96 and was protective in most of the study region with small pocket of increased odds **(*Figure 4.12*)**. One significant cluster of high odds (O/E = 46.90) was found in the southeastern-most region, just below Lake Victoria. Distance to cluster center had an odds ratio range of 0.75 – 1.68 with a small band of high odds of malaria infection by distance crossing from the southeast to the northwest **(*Figure 4.13*)**. Three significant clusters of ORs (O/E = 1.39 – 1.08) appeared in this band, and two clusters of low ORs were indicated in the westernmost area (O/E = 0.91) and the far northeast (O/E = 0.87). Whether a household had open eaves displayed an odds ratio range of 0.13 – 3.08 **(*Figure 4.14*)**. A line of four significant high OR clusters (O/E = 2.89 – 1.73) crosses the center of the map from north to south. Six low OR clusters were indicated, two in the southeast along the shore of Lake Victoria (O/E = 0.45 – 0.46), three in the center of the map (O/E = 0.33 – 0.61), and one in the far northeast (O/E = 0.39). Type of roofing material had an odds ratio range of 0.02 – 6.86 with a fairly scattered distribution of high and low odds of malaria infection with roofing material type **(*Figure 4.15*)**. Four high OR clusters (O/E = 4.26 – 1.59) sit in the center of the map, with two low OR clusters in the southeast (O/E = 0.59) and northwest (O/E = 0.55). Head of household schooling had an odds ratio range of 0.06 – 4.71 with high odds ratios along the eastern regions of the study area

(*Figure 4.16*). Four high OR clusters (O/E = 2.97 – 1.62) were found along the eastern side of the study area, with two low OR clusters in the far northwest (O/E = 0.69) and southwest (O/E = 0.45). Household wealth had an odds ratio range of 0.31 – 2.34 with high odds ratios mostly concentrated in the southern region of the study area **(*Figure 4.17*)**. Two high OR clusters (O/E = 1.74, 1.35) were indicated in the southern parts of the study region, and one low OR cluster (O/E 0.77) dominates the center of the map. ITN or LLIN use had an odds ratio range of 0.05 – 26.22 with ITNs or LLINs being mostly protective or non-significant across the study area **(*Figure 4.18*)**. One high odds cluster was indicated in the center-north of the study area (O/E = 3.85) and one cluster of extremely high odds was be found just to the north (O/E = 13.57) although it covered only a very small area.

4.5 Discussion

This study contributes further evidence that malaria is spatially heterogeneous even in localized regions. This finding is consistent with similar studies on malaria in Tanzania (West, et al., 2013), Senegal (Ndiath, et al., 2015), Kenya (Homan, et al., 2016), Zimbabwe (Manyangadze, Chimbari, Macherera, & Mukaratirwa, 2017), and Indonesia (Hasyim, et al., 2018). In this study the geographically-weighted adjusted odds of malaria infection crossed "1" for each variable under consideration, provided odds ratio ranges of up to almost 30 (***Table 4.4***), and provided substantially better model fit (AICc 3,133 vs 4,296). The SaTScan analyses further demonstrate that there are spatial clusters of greater and lower than expected odds of malaria as predicted by each variable, which can be used to inform targeted interventions (Bousema, et al., 2010).

Increasing odds of infection by *Plasmodium spp.* as predicted by age is in line with recent studies in western Kenya (Were, et al., 2018), Tanzania (Froeschl, et al., 2018), and Vietnam (Bannister-Tyrrell, et al., 2018). While this association does not necessarily track with malaria cases, as passive immunity is developed with age (Marsh, 1992), infected asymptomatic individuals are still reservoirs for endemic transmission (Bousema, Okell, Felger, & Drakeley, 2014; Okell, et al., 2012). In our study, we found locally concentrated odds of malaria as predicted by age in an area where age was predominately lower than expected.

Increased elevation was found to be a protective factor for malaria infection in the present study as has been shown in multiple studies previously in Uganda (Roberts & Matthews, 2016), Cameroon (Tonye, Kouambeng, Wounang, & Vounatsou, 2018), and Tanzania, where an extensive study in the Usambara Mountains explored the elevation-temperature-rainfall interaction on *Anopheles spp.* vector density and *P. falciparum* infection and found a direct correlation with elevation (Bødker, et al., 2003). The extremely large range of odds for malaria infection as predicted by elevation (ORs = 0.00 – 29.96) in the Muleba area of Tanzania are likely indicative of the dramatic range of elevation in the region, with households ranging from 1,075 – 1,654 meters above sea level. Households at high altitude may find themselves above the range at which *Anopheles spp.* are able to survive and are therefore at extremely low risk of exposure to *P. falciparum.* Low-lying areas may be at extremely high risk of malaria infection due to the increased likelihood of being adjacent to perpetually standing water. Large odds of malaria by increasing elevation are likely statistical anomalies, where small increases in elevation in extreme low-lying areas are associated with malaria infection due to unmeasured externalities. These areas should be carefully examined for other effects that could be causing a local increase in malaria infection.

Open eaves are significantly associated with odds of malaria infection as has been document in previous studies in Malawi (Walldorf, et al., 2015) and Equatorial Guinea (Bradley, et al., 2013). Natural roofing material was non-significant in the adjusted model (p=0.0603), which is unusual as studies often find associations between natural or unfinished roofs or walls and malaria infection (Ghebreyesus, et al., 2000; Homenauth, Kajeguka, & Kulkarni, 2017; Roberts & Matthews, 2016), or the density of *Anopheles spp.* as in Malawi (Dear, et al., 2018), which could provide a mediating association with malaria infection. These housing quality measures both provide a potential point-of-access for the *Anopheles* mosquito – and therefore potentially the *Plasmodium* parasite – that would not be available otherwise. It is however, not unheard of for these factors to be non-significant when controlling for socioeconomic status (Brooker, et al., 2004; Bousema, et al., 2010).

As in this study, socioeconomic status by wealth index has been repeatedly shown to be a strong predictor of malaria infection and cases with results from Kenya (Were, et al., 2018), Cameroon (Tonye, Kouambeng, Wounang, & Vounatsou, 2018), and Uganda (Ssempiira, et al., 2018) just in the previous year, as well as previous results from the same study area in Tanzania (West, et al., 2013). Like with most diseases it shares an association with, poverty exacerbates the conditions and circumstances that allow malaria to thrive. Similarly, as in this analysis, head of household education has shown an association with malaria in multiple recent studies in Uganda (Roberts & Matthews, 2016), Cameroon (Tonye, Kouambeng, Wounang, & Vounatsou, 2018). Higher levels of education are associated with knowledge of malarial transmission modes (Gessler, et al., 1995; Mboera, et al., 2010) which can ensure that interventions are adequately sought after and applied. Head of household education in particular could contribute to ensuring that children are inside and away from vectors during biting hours.

ITN/LLIN use was concentrated in the southern regions of the study area where it appears to have a reasonably consistent protective effect as expected. Household-level usage of preventive measures is generally associated with reductions in malaria prevalence, although these associations may be confounded by several factors. The use of ITNs and LLINs have been demonstrated to reduce the risk of malaria infection (Lengeler, 2009), but ownership does not directly translate to use. Proper LLIN use requires that the net must be hung over a bed or other sleeping space and that individuals within the household must sleep under it each night, which may be impacted by household education (Atieli, et al., 2011). The impact of LLINs may be mitigated due to confounding by malaria endemicity as there is an association between malaria prevalence and both net ownership and usage (Burgert, Bradley, Arnold, & Ekert, 2014; Atieli, et al., 2011).

4.5.1. Information and Insecticide Resistance

The central region of the study is characterized by a band of higher than expected malaria infection that overlaps directly with several covariates of higher than expected odds of infection. In particular, even when controlling for other covariates and using local effects, the odds of household eave type, distance to cluster center, roofing type, and head of household-associated malaria infection was highest in the center region of the study area. Additionally, LLIN/ITN use was found to be dramatically predictive of malaria infection in the northern region (up to 13.57 times more predictive than the rest of the study area), where bednet use is clearly less common despite being at high risk for malaria. The fact that there is a disparity in use despite being a higher SES region could suggest that there is a failure in dissemination of information regarding

the benefits of anti-malarial interventions as has been documented previously (Heggenhougen, Hackethal, & Vivek, 2003).

An additional, and not entirely separate possibility is that there is a pocket of increased insecticide resistance in this region. Tanzania is one of several countries in the African region that is experiencing dramatic insecticide resistance, particularly with reference to the four main WHO recommended insecticides (WHO, 2017). A WHO report in 2016 notes that there was a heterogenous trend to malaria resistance across 5 countries in sub-Saharan Africa that exhibited fine scale variations (WHO, 2016; Kleinschmidt, et al., 2015). As information regarding novel insecticides and their use is still being disseminated amongst academics and policy-makers (Protopopoff, et al., 2018; WHO, 2017) is quite likely that the fact that IRS is both concentrated and ineffective in the region may be the result of spraying with an insecticide that is inhibited by resistance.

4.5.2. Targeting Interventions in Heterogeneous Areas

In one instance, increased odds of malaria infection predicted by age (Age OR Cluster 4) overlapped with a low age cluster (Age Cluster 2) with moderate LLIN/ITN use and efficacy. Given the increased vulnerability of young children to severe malaria outcomes (WHO, 2017), this presents an opportunity for the introduction of a targeted intervention. As there is no indication that either intervention would have decreased efficacy in the area, the distribution of LLIN/ITNs to older children and young adults could provide targeted and direct impact on the odds of malaria in that area.

It should be noted that an attempt at hotspot-targeted control has been made previously to some small effect in the Kenyan highlands (Bousema, et al., 2016). This intervention was

conducted under the assumption that targeted control could have impact on greater areas than the intervention itself, which was not observed. While the effect on the hotspot itself was modest (10.2% reduction in parasite prevalence), this suggests that targeted interventions should be made with the understanding that their impacts may not extend beyond the targeted areas and should not be exclusively used in place of larger-scale interventions where malaria is endemic.

4.5.3. Limitations

This study was completed using an existing dataset and it was therefore not possible to include other variables often associated with malaria infection. The inclusion of entomological and environmental data associated with malaria transmission intensity may have improved the model, however these data were not within the scope of the current study. Nonetheless, elevation was considered as a variable in the model, since altitude correlates well with *Anopheles spp.* density (Protopopoff, et al., 2009; Kelly-Hope, Hemingway, & McKenzie, 2009; Daygena, Massebo, & Lindtjørn, 2017) and temperature (Goody & Walker, 1972).

Additionally, because this study was conducted on a single cross-sectional it is impossible to assess the temporality of risk factors on the outcome. With the possible exception of the malaria interventions (ITN/LLIN and IRS) it is unlikely that childhood malaria infection would have any bearing on the other covariates under study.

4.6 Conclusion

Using data from a large cross-sectional survey in Muleba, Tanzania, this study demonstrates an enormous variability in the odds of malaria infection in children by multiple predictor variables. The GWLR model provides greater model fit than a global model developed

using standard multivariable logistic regression, implying that the underlying spatial variation in predictor variables is important to consider when exploring spatial variations in malaria prevalence. In particular, age of child and several household factors, such as elevation, type of eaves, head of household education, wealth quintile, and household bednet use were all found to be associated with malaria in the adjusted model and all variables were found to have significant spatial trends. Future malaria studies should ensure that the spatial nature of malaria is carefully considered in their analytic approach.

4.7 References

APHA. (2015). *Control of Communicable Diseases Manual* (20th ed.). (D. L. Heymann, Ed.) Washington: Alpha Press.

ArcGIS. (2015). ArcGIS 10.4.1. Redlands, CA, USA: Envicornmental Systems Research Institute (ESRI) Inc.

Atieli, H. E., Zhou, G., Afrane, Y., Lee, M., Mwanzo, I., Githeko, A. K., & Yan, G. (2011). Insecticide-treated net (ITN) ownership, usage, and malaria transmission in the highlands of western Kenya. *Parasites & Vectors, 4*(113), 1-10.

Bannister-Tyrrell, M., Xa, N. X., Kattenberg, J. H., Van, N. V., Dung, V. K., Hieu, T. M., . . . Erhart, A. (2018). Micro-epidemiology of malaria in an elimination setting in Central Vietnam. *Malaria Journal, 17*(119), 1-15.

Bødker, R., Akida, J., Shayo, D., Kisinza, W., Msangeni, H. A., Pedersen, E. M., & Lindsay, S. W. (2003). Relationship Between Altitude and Intensity of Malaria Transmission in the Usambara Mountains, Tanzania. *Journal of Medical Entomology, 40*(5), 706-717.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8*, 3-62.

Bousema, T., Drakeley, C., Gesase, S., Hashim, R., Magesa, S., Mosha, F., . . . Gosling, R. (2010). Identification of Hot Spots of Malaria Transmission for Targeted Malaria Control. *The Journal of Infectious Diseases, 201*(11), 1764-1774.

Bousema, T., Okell, L., Felger, I., & Drakeley, C. (2014). Asymptomatic malaria infections: detectability, transmissibility and public health relevance. *Nature Reviews Microbiology, 12*(12), 833-840.

Bousema, T., Stresman, G., Baidjoe, A. Y., Bradley, J., Knight, P., Stone, W., . . . Cox, J. (2016). The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLOS Medicine, 13*(4).

Bradley, J., Rehman, A. M., Schwabe, C., Vargas, D., Monti, F., Ela, C., . . . Kleinschmidt, I. (2013). Reduced Prevalence of Malaria Infection in Children Living in Houses with Window Screening or Closed Eaves on Bioko Island, Equatorial Guinea. *PLoS ONE, 8*(11), 1-7.

Brooker, S., Clarke, S., Njagi, J. K., Polack, S., Mugo, B., Estamblale, B., . . . Cox, J. (2004). Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Tropical Medicine & International Health, 9*(7), 757-766.

Burgert, C. R., Bradley, S. E., Arnold, F., & Ekert, E. (2014). Improving estimates of insecticide-treated mosquito net coverage from household surveys: Using geographic coordinates to account for endemicity. *Malaria Journal, 13*(254), 1-13.

Daygena, T. Y., Massebo, F., & Lindtjørn, B. (2017). Variation in species composition and infection rates of Anopheles mosquitoes at different altitudinal transects, and the risk of malaria in the highland of Dirashe Woreda, south Ethiopia. *Parasites & Vectors, 10*(343), 1-13.

Dear, N. F., Kadangwe, C., Mzilahowa, T., Bauleni, A., Mathanga, D. P., Duster, C., . . . Wilson, M. L. (2018). Household-level and surrounding peri-domestic environmental characteristics associated with malaria vectors Anopheles arabiensis and Anopheles funestus along an urban–rural continuum in Blantyre, Malawi. *Malaria Journal, 17*(229), 1-14.

Froeschl, G., Saathoff, E., Kroidl, I., Berens-Risha, N., Clowes, P., Maboko, L., . . . Hoelscher, M. (2018). Reduction of malaria prevalence after introduction of artemisinin-combination-therapy in Mbeya Region, Tanzania: results from a cohort study with 6773 participants. *Malaria Journal, 17*(245), 1-11.

Gessler, M. C., Msuya, D. E., Nkunya, M. H., Schär, A., Heinrich, M., & Tanner, M. (1995). Traditional healers in Tanzania: the perception of malaria and its causes. *Journal of Ethnopharmacology, 48*, 119-130.

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R., Johnston, G. L., . . . Hay, S. I. (2011). A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria Journal, 10*(378), 1-16.

Ghebreyesus, T. A., Haile, M., Witten, K. H., Getachew, A., Yohannes, M., Lindsay, S. W., & Byass, P. (2000). Household risk factors for malaria among children in the Ethiopian highlands. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 94*(1), 17-21.

Goody, R. M., & Walker, J. C. (1972). Atmospheric Temperatures. In R. M. Goody, & J. C. Walker, *Atmospheres* (pp. 44-70). Prentice Hall.

GWR4. (2015, 03). GWR4 v4.0.9. Maynooth , Ireland: National Center for Geocomputation, National University of Ireland Maynooth and Department of Geography, Ritsumeikan University, Japan.

Hasyim, H., Nursafingi, A., Haque, U., Montag, D., Groneberg, D. A., Dhimal, M., . . . Müller, R. (2018). Spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia. *Malaria Journal, 17*(87), 1-15.

Heggenhougen, H. K., Hackethal, V., & Vivek, P. (2003). *The behavioural and social aspects of malaria and its control: An introduction and annotated bibliography.* UNDP/World Bank/WHO, Special Programme for Research and Training in Tropical Diseases. Geneva: World Health Organization.

Homan, T., Marie, N., Hiscox, A., Di Pasquale, A., Kiche, I., Onoka, K., . . . Takken, W. (2016). Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: an explorative study. *Malaria Journal, 15*.

Homenauth, E., Kajeguka, D., & Kulkarni, M. A. (2017). Principal component analysis of socioeconomic factors and their association with malaria and arbovirus risk in Tanzania: a sensitivity analysis. *Socioeconomic Factors and Health, 71*, 1046-1051.

Jung, I., Kulldorff, M., & Klassen, A. (2007). A spatial scan statistic for ordinal data. *Statistics in Medicine, 26*, 1594-1607.

Kelly-Hope, L. A., Hemingway, J., & McKenzie, F. E. (2009). Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya. *Malaria Journal, 8*(268), 1-8.

Kleinschmidt, I., Mnzava, A. P., Kafy, H. T., Mbogo, C., Bashir, A. I., Bigoga, J., . . . Donnelly, M. J. (2015). Design of a study to determine the impact of insecticide resistance on malaria vector control: a multi-country investigation. *Malaria Journal, 14*(282), 1-13.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods, 26*, 1481-1496.

Kulldorff, M., Huang, L., & Konty, K. (2009). A spatial scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics, 8*(58), 1-9.

Lengeler, C. (2009). Insecticide-treated bed nets and curtains for perventing malaria. (C. I. Group, Ed.) *Cochrane Database of Systematic Reviews*(2), pp. 1-55.

Manyangadze, T., Chimbari, M. J., Macherera, M., & Mukaratirwa, S. (2017). Micro-spatial distribution of malaria cases and control strategies at ward level in Gwanda district, Matabeleland South, Zimbabwe. *Malaria Journal, 16*(476), 1-11.

Marsh, K. (1992). Malaria - a neglected disease? *Parasitology, 104*, S53-S69.

Mboera, L. E., Shayo, E. H., Senkoro, K. P., Rumisha, S. F., Mlozi, M. R., & Mayala, B. K. (2010). Knowledge, perceptions and practices of farming communities on linkages between malaria and agriculture in Mvomero District, Tanzania. *Acta Tropica, 113*, 139-144.

Midekisa, A., Beyene, B., Mihretie, A., Bayabil, E., & Wimberly, M. C. (2015). Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. *Parasites & Vectors, 8*(339), 1-11.

National Bureau of Statistics. (2016). *Basic Demographic and Socio-Economic Profile Kagera Region.* Ministry of Finance, National Bureau of Statistics. Dar es Salaam: The United Republic of Tanzania.

Ndiath, M. M., Cisse, B., Ndiaye, J. L., Gomis, J. F., Bathiery, O., Dia, A. T., . . . Faye, B. (2015). Application of geographically-weighted regression analysis to assess risk factors for malaria hotspots in Keur Soce health and demographic surveillance site. *Malaria Journal, 14*(463).

Okell, L. C., Bousema, T., Griffin, J. T., Oue´draogo, A. L., Ghani, A. C., & Drakeley, C. J. (2012). Factors determining the occurence of submicroscopic malaria infections and their relevance for control. *Nature Communications, 3*(1237), 1-9.

Protopopoff, N., Bortel, W. V., Speybroeck, N., Geertruyden, J. V., Baza, D., D'Alessandro, U., & Coosemans, M. (2009). Ranking Malaria Risk Factors to Guide Malaria Control Efforts in African Highlands. *PLoS One, 4*(11).

Protopopoff, N., Mosha, J. F., Lukaloe, E., Charlwood, J. D., Wright, A., Mwalimu, C. D., . . . Rowland, M. (2018). Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two fact. *Lancet, 391*, 1577-88.

Roberts, D., & Matthews, G. (2016). Risk factors of malaria in children under the age of five years old in Uganda. *Malaria Journal*(15).

SAS. (2012). SAS v9.4. Cary, NC, USA: SAS Institute Inc.

SaTScan. (2016, 08). SaTScan v9.4.4. Calverton, Maryland: Harvard Medical School, Boston and Information Management Services, Inc. Retrieved from http://www.satscan.org/

Sena, L., Deressa, W., & Ali, A. (2015). Correlation of climate variability and malaria: A retrospective comparative study, southwest Ethiopia. *Ethiopian Journal of Health Science, 25*(2), 129-138.

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., . . . Hay, S. I. (2012). A global map of dominant malaria vectors. *Parasites & Vectors, 5*(69), 1-11.

Ssempiira, J., Nambuusi, B., Kyozira, C., Rutazaana, D., Mukooyo, E., Opigo, J., . . . Vounatsou, P. (2018). The efect of case management and vector-control interventions on space–time patterns of malaria incidence in Uganda. *Malaria Journal, 17*(162), 1-11.

Thomson, M. C., Mason, S. J., Phindela, T., & Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. *Journal of Tropical Medicine & Hygiene, 73*(1), 214-221.

Tonye, S. G., Kouambeng, C., Wounang, R., & Vounatsou, P. (2018). Challenges of DHS and MIS to capture the entire pattern of malaria parasite risk and intervention efects in countries with diferent ecological zones: the case of Cameroon. *Malaria Journal, 17*(156), 1-14.

Walldorf, J. A., Cohee, L. M., Coalson, J. E., Bauleni, A., Nkanauena, K., Kapito-Tembo, A., . . . Laufer, M. K. (2015). School-age children are a resevoir of malaria infection in Malawi. *PLoS ONE, 10*(7), 1-13.

Were, V., Buff, A. M., Desai, M., Kariuki, S., Samuels, A., ter Kuile, F. O., . . . Niessen, L. (2018). Socioeconomic health inequality in malaria indicators in rural western Kenya: evidence from a household malaria survey on burden and care-seeking behaviour. *Malaria Journal, 17*(166), 1-10.

West, P. A., Protopopoff, N., Rowland, M., Cumming, E., Rand, A., Drakeley, C., . . . Kleinschmidt, I. (2013). Malaria Risk Factors in North West Tanzania: The Effect of Spraying, Nets and Wealth. *PLOS One, 8*(6), 1-12.

WHO. (2016). *Implications of insecticide resistance for malaria vector control.* Global Malaria Programme. Geneva: World Health Organization.

WHO. (2017). *Conditions for deployment of mosquito nets treated with a pyrethroid and piperonyl butoxide.* Global Malaria Programme. Geneva: World Health Organization.

WHO. (2017). *World Malaria Report.* Geneva: World Health Organization.

WHO. (2018). *World Malaria Report 2018.* Geneva: World Health Organization.

Wiebe, A., Longbottom, J., Gleave, K., Shearer, F. M., Sinka, M. E., Massey, N. C., . . . Moyes, C. L. (2017). Geographical distributions of African malaria vector sibling species and evidence for insecticide resistance. *Malaria Journal, 16*(85), 1-10.

Table 4.1 *Characteristics of Baseline Clinical Population by Rapid Diagnostic Test Result*

Parameter	Negative RDT n = 1296 μ (95%CI) or n (%)	Positive RDT n = 2246 μ (95%CI) or n (%)	p-value*
Age (years)	6.07 (5.85-6.28)	6.75 (6.59 -6.91)	<0.0001
Elevation (meters)	1345 (1340-1351)	1297 (1294-1301)	<0.0001
Distance from Cluster Center (meters)	921 (887-956)	988 (960-1015)	0.0035
Household has Open Eaves			
Yes	695 (32.09)	1471 (67.91)	<0.0001
No	601 (43.68)	775 (56.32)	
Roofing Material			
Grass/Leaves & Partial Metal	167 (25.30)	493 (74.70)	<0.0001
Metal Sheets	1129 (39.17)	1753 (60.83)	
Head of Household Schooling			
No Education	285 (29.75)	673 (70.25)	
Primary Education	954 (38.58)	1519 (61.42)	<0.0001
Secondary/Tech or Higher Education	57 (51.35)	54 (48.65)	
Wealth Quintile			
Poorest	180 (28.85)	444 (71.15)	
Poorer	209 (29.65)	496 (70.35)	
Middle	280 (37.84)	460 (62.16)	<0.0001
Less Poor	290 (40.73)	422 (59.27)	
Least Poor	337 (44.28)	424 (55.72)	
ITN/LLIN Used			
Yes	736 (34.75)	1382 (65.25)	0.0056
No	560 (39.33)	864 (60.67)	
Indoor Residual Spraying			
Yes	183 (34.86)	342 (65.14)	0.3719
No	1113 (36.92)	1902 (63.08)	

*p-values are calculated using t-test or chi-square as appropriate

Table 4.2 *Significant Clusters of Predictors and Measures of Spatial Autocorrelation by Moran's I at Baseline*

Parameter	Cluster	Observed*	Expected*	O/E	p value**	Moran's I Fit	Moran's I p value
Malaria Infection						0.2460	<0.0001
High Risk Clusters	6	122	80.52	1.52	7.54E-14		
	5	134	89.39	1.50	3.36E-14		
	4	149	100.81	1.48	2.00E-14		
	7	130	88.13	1.48	4.94E-12		
	1	535	410.83	1.30	1.00E-17		
Low Risk Clusters	2	114	211.76	0.54	1.00E-17		
	10	48	89.39	0.54	4.82E-08		
	9	48	91.30	0.53	8.57E-09		
	11	18	47.55	0.38	1.59E-07		
	3	31	89.39	0.35	1.00E-17		
	12	11	33.60	0.33	7.64E-06		
	8	6	32.97	0.18	5.92E-10		
Age						0.0253	<0.0001
High Age Cluster	1	7.66	6.28	1.22	0.002		
Low Age Cluster	2	5.09	6.43	0.79	0.002		
Elevation						0.9668	<0.0001
High Elevation Clusters	4	15.72	13.12	1.20	0.001		
	1	14.31	12.97	1.10	0.001		
	3	13.87	12.99	1.07	0.001		
Low Elevation Clusters	8	12.75	13.19	0.97	0.001		
	6	12.41	13.19	0.94	0.001		
	7	12.35	13.18	0.94	0.001		
	5	12.24	13.19	0.93	0.001		
	2	11.96	13.26	0.90	0.001		
Distance from Cluster Center						0.8600	<0.0001
Rural Clusters	9	38.84	9.60	4.05	0.001		
	2	26.70	9.48	2.82	0.001		
	6	25.15	9.54	2.64	0.001		
	7	21.98	9.54	2.30	0.001		
	5	20.15	9.49	2.12	0.001		
	4	16.57	9.35	1.77	0.001		
	3	16.04	9.29	1.73	0.001		
Urban Clusters	10	6.73	9.89	0.68	0.001		
	1	6.40	10.56	0.61	0.001		
	8	2.47	9.77	0.25	0.001		
Household has Open Eaves						0.0979	<0.0001
High Open Eave Clusters	3	28	10.88	2.57	2.47E-08		
	1	187	117.71	1.59	7.94E-12		
Low Open Eave Clusters	2	8	41.18	0.19	6.67E-09		
	4	0	20.59	0.00	3.35E-08		
Roof						0.1420	<0.0001
High Leaf Roof Clusters	3	17	3.73	4.56	7.52E-06		
	1	57	19.94	2.86	1.25E-10		
High Metal Roof Cluster	2	46	95.4	0.48	5.43E-06		
Head of Household Schooling						0.1070	<0.0001

High Education Clusters	10	4	0.16	25.00	0.049		
	6	6	0.72	8.33	0.005		
	4	5	1.50	3.33	0.002		
	2	36	16.64	2.16	0.001		
	3	61	44.50	1.37	0.001		
	9	41	29.91	1.37	0.014		
Low Education Clusters	8	10	2.70	3.70	0.012		
	5	12	3.25	3.69	0.003		
	7	30	13.25	2.26	0.008		
	1	189	113.06	1.67	0.001		
Wealth Quintile						0.0733	<0.0001
High Wealth Clusters	2	14	3.01	4.65	0.001		
	3	11	2.36	4.65	0.004		
	5	18	7.3	2.46	0.001		
	6	15	7.52	1.99	0.003		
	7	25	15.04	1.66	0.002		
	8	119	73.26	1.62	0.001		
Low Wealth Clusters	1	35	19.56	1.79	0.001		
	4	60	38.58	1.56	0.001		
ITN/LLIN Used						0.1410	<0.0001
High Bednet Use Clusters	2	60	25.33	2.37	8.90E-16		
	5	57	29.75	1.92	3.15E-06		
	1	316	210.26	1.50	1.00E-17		
Low Bednet Use Clusters	4	90	163.63	0.55	3.25E-11		
	3	103	191.77	0.54	6.88E-15		
Indoor Residual Spraying						0.0991	<0.0001
High IRS Cluster	1	1303	1224.01	1.07	1.55E-09		
Low IRS Clusters	2	70	104.77	0.67	2.74E-09		
	3	34	60.48	0.56	1.59E-08		

*Observed and Expected show mean average in cluster and out of cluster respectively for continuous data
**Dichotomous data is cluster-analysed with the SaTScan Bernoulli model which does not account for multiple testing. A Bonferroni correction is applied for a significant value of 2.65e-5. For all other cases clusters below $p>0.05$ are discarded.

Figure 4.1 *Spatial distribution and clustering of malaria infection in baseline population*

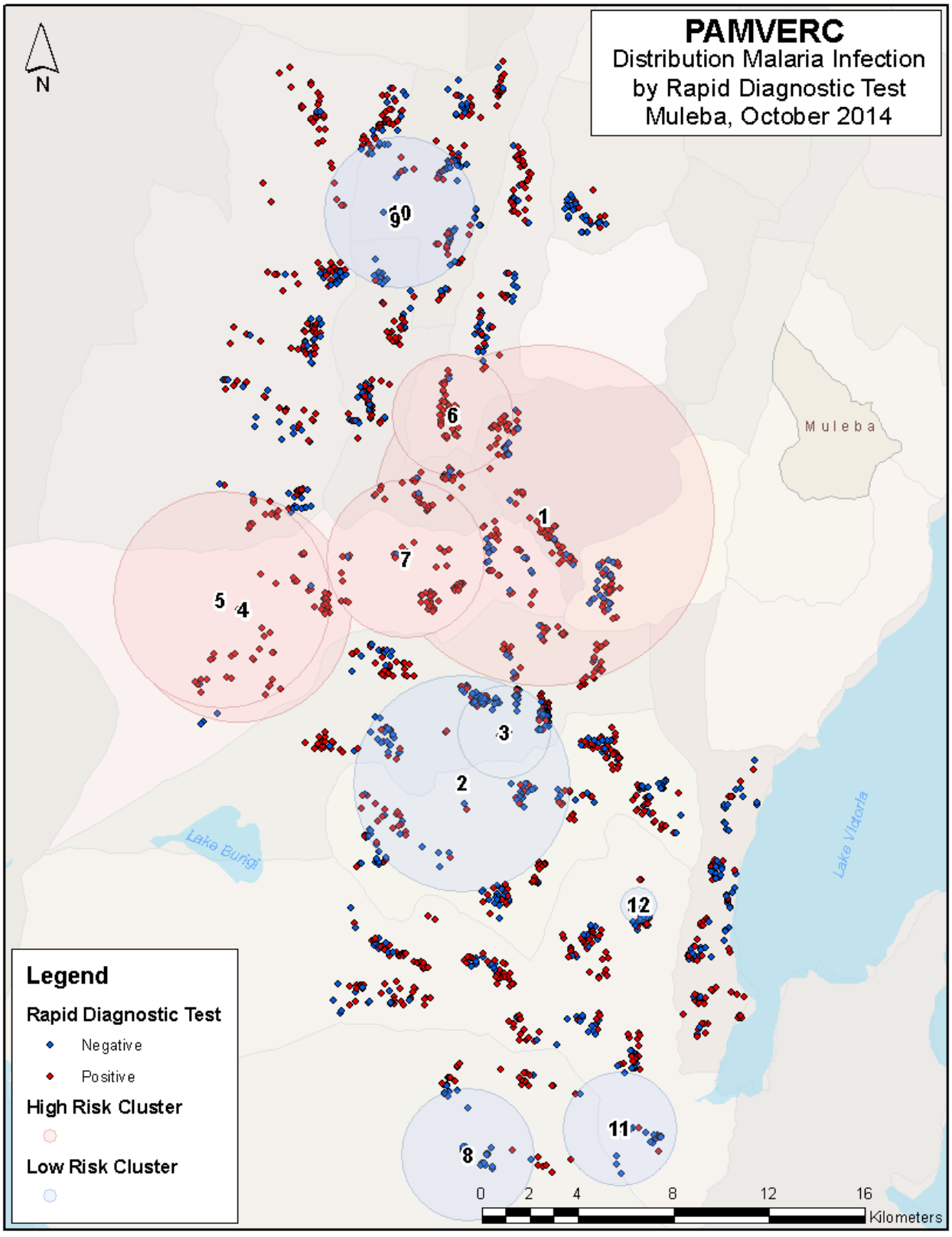


Figure 4.2 *Spatial distribution and clustering of age in baseline population*

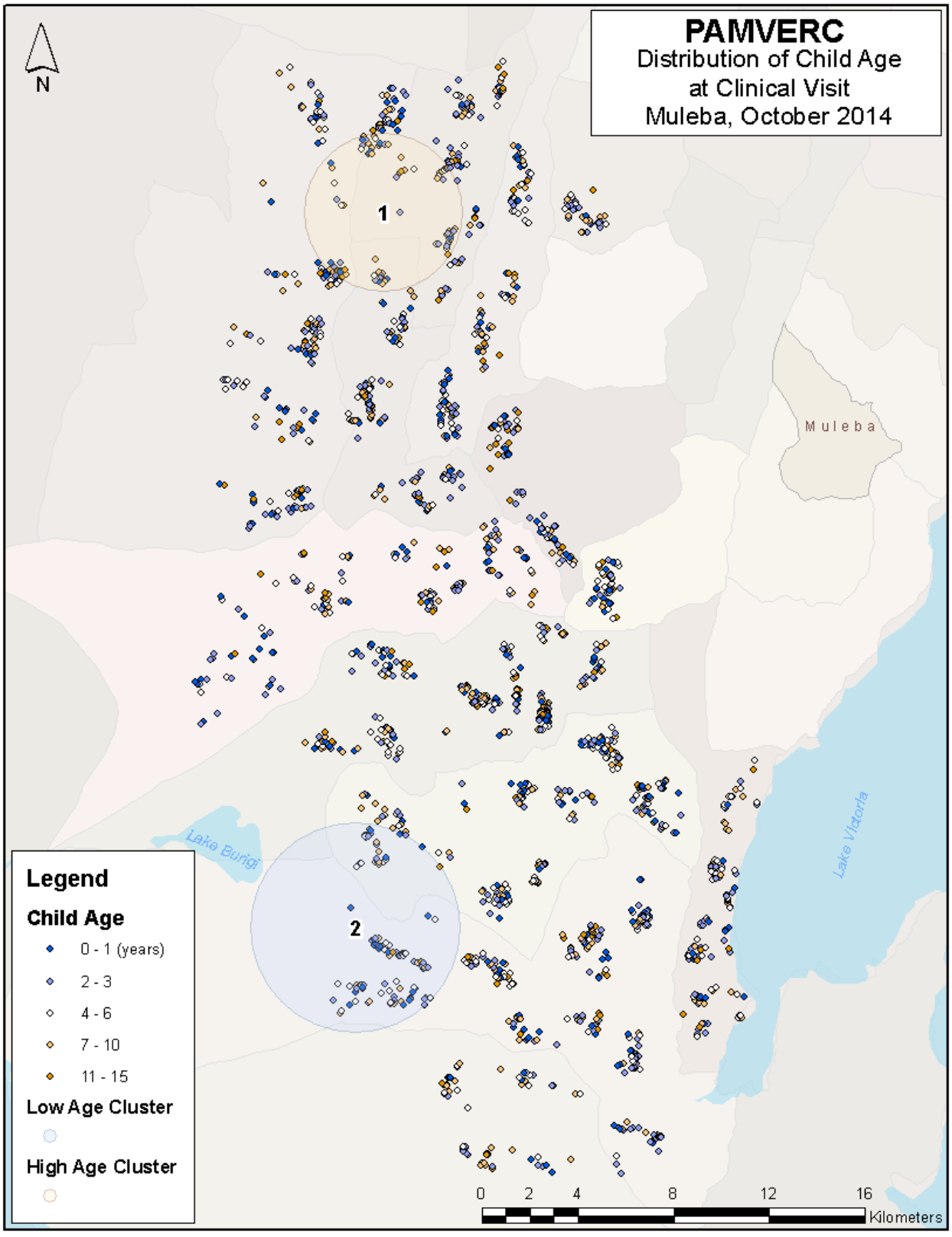


Figure 4.3 *Spatial distribution and clustering of elevation in baseline population*

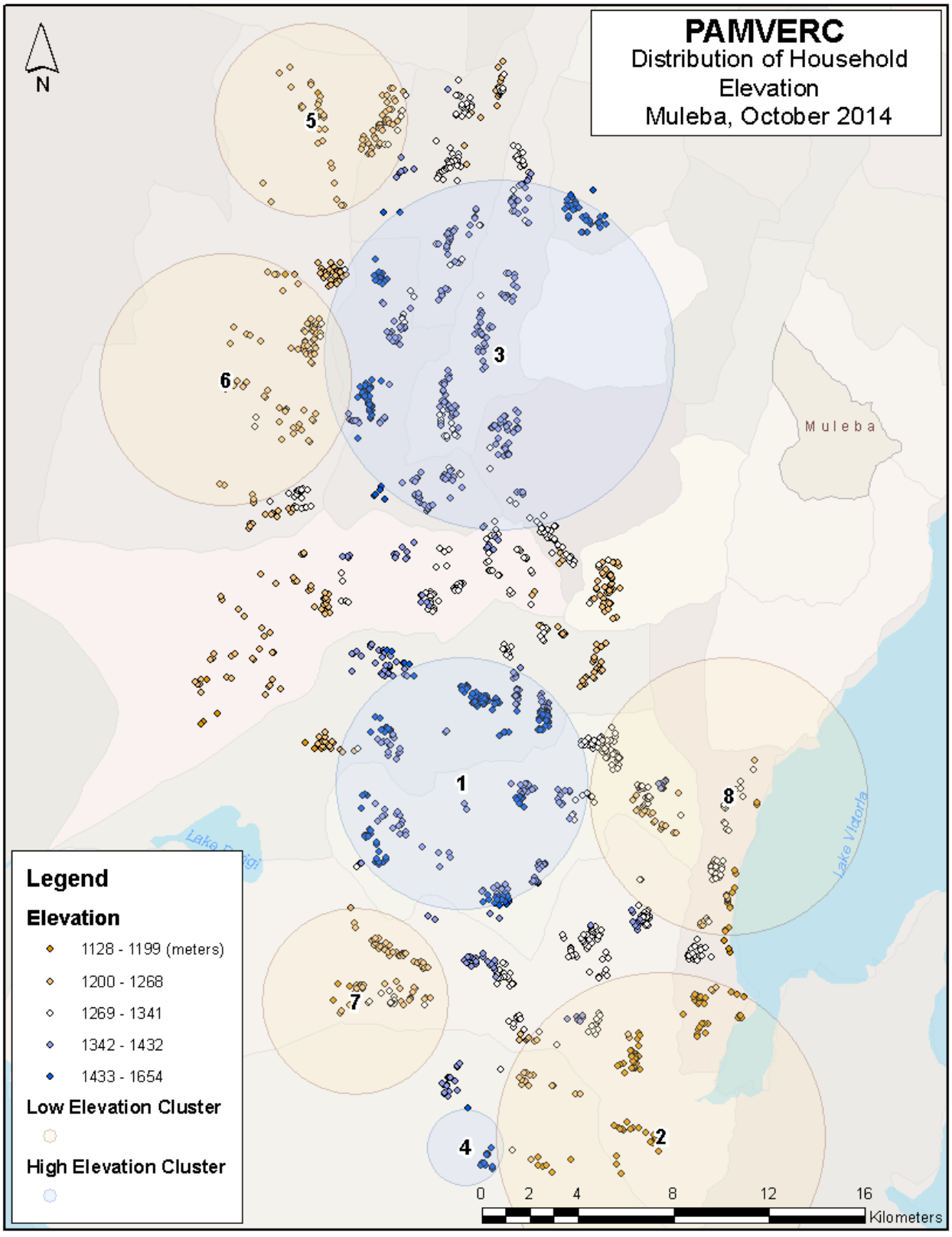


Figure 4.4 *Spatial distribution and clustering of distance to cluster center in baseline population*

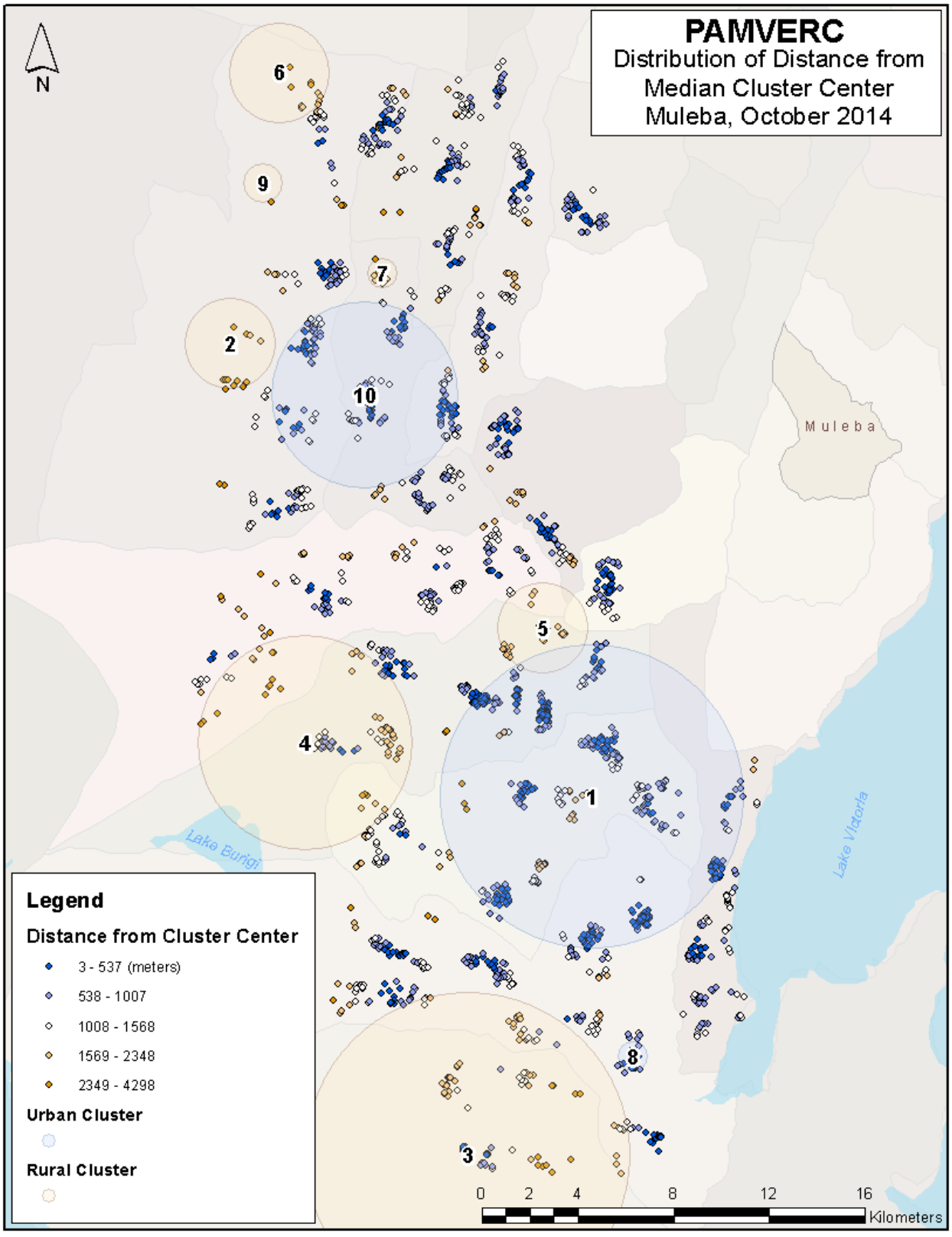


Figure 4.5 *Spatial distribution and clustering of type of eaves in baseline population*

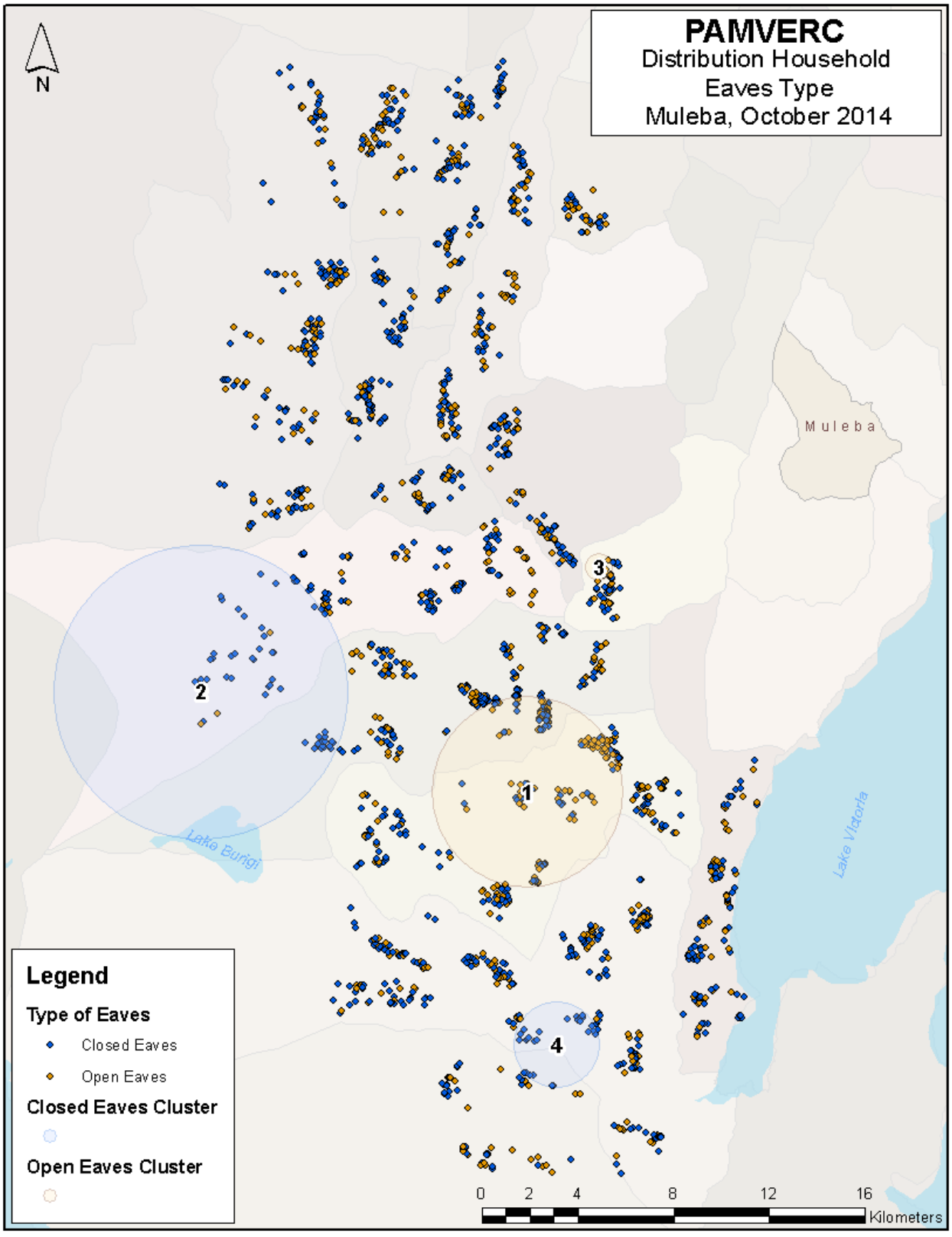


Figure 4.6 *Spatial distribution and clustering of roofing material in baseline population*

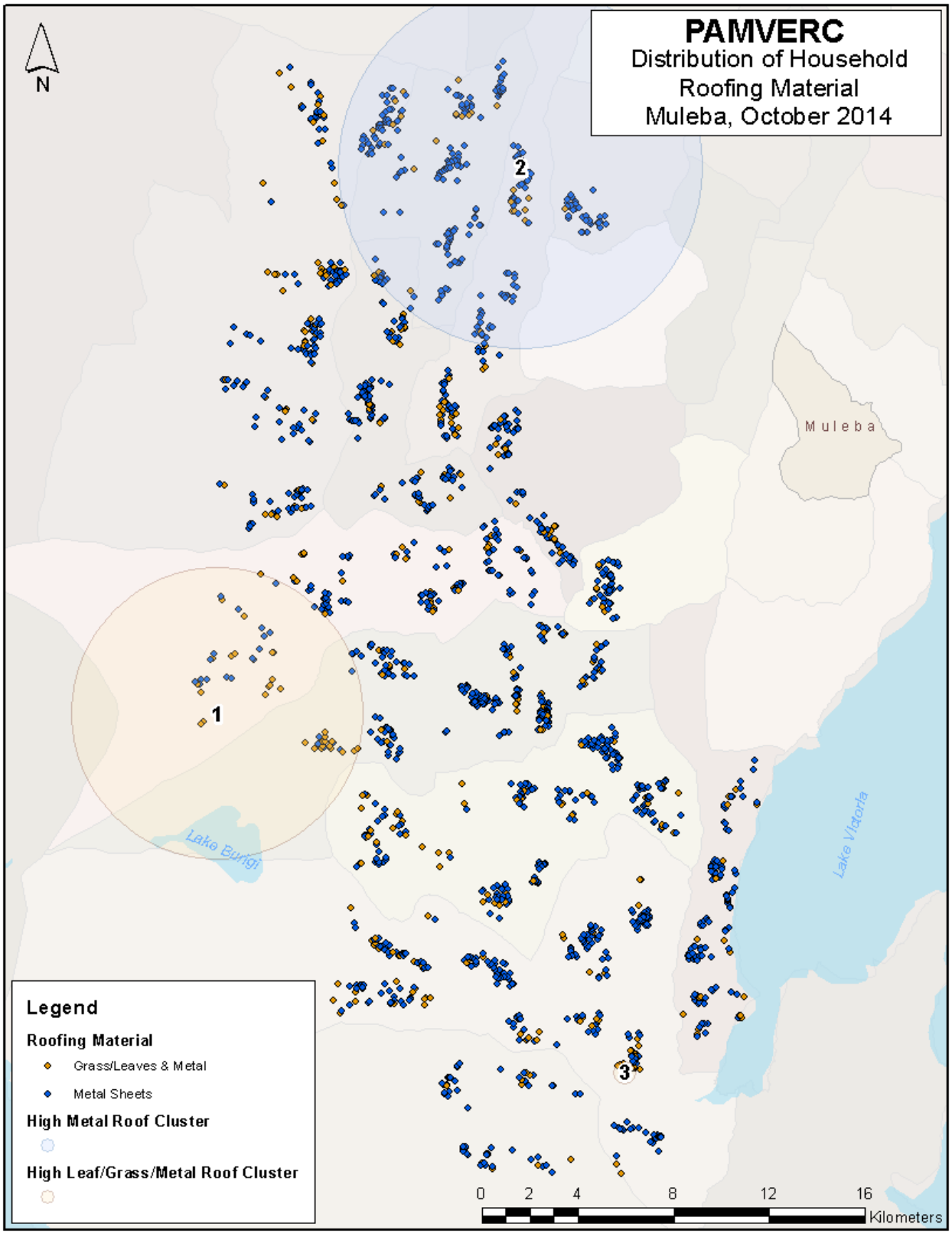


Figure 4.7 *Spatial distribution and clustering of head of household schooling in baseline population*

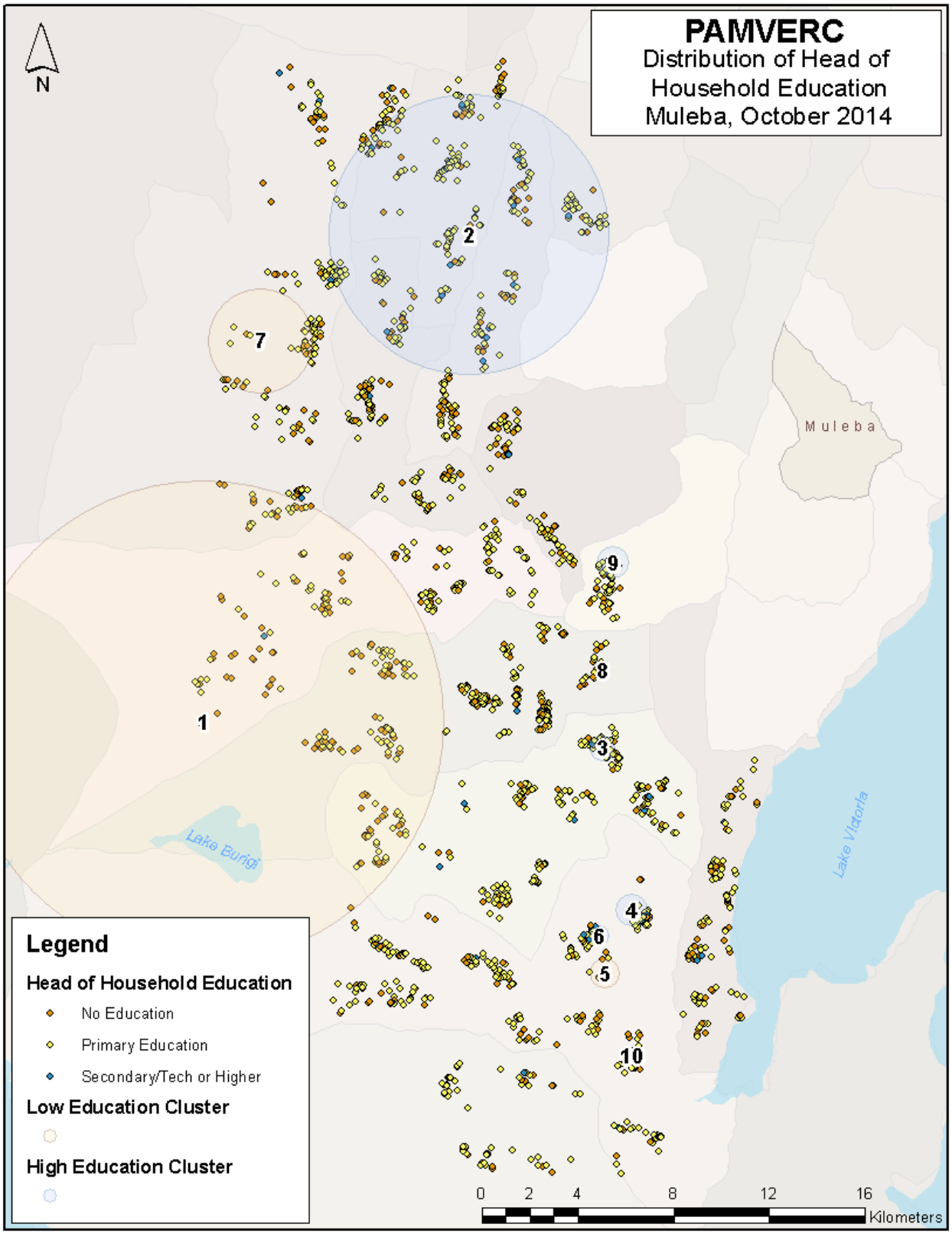


Figure 4.8 *Spatial distribution and clustering of wealth quintiles in baseline population*

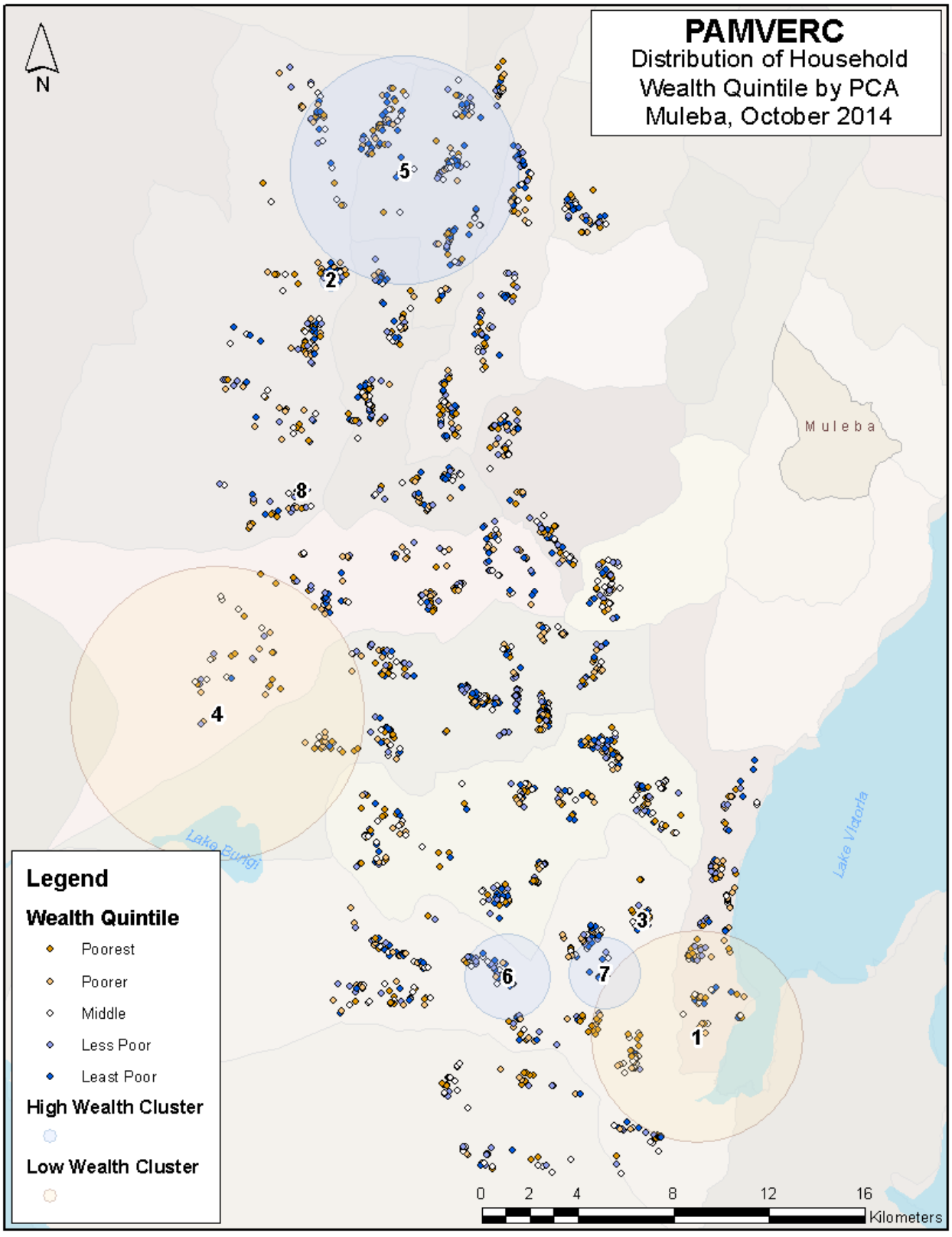


Figure 4.9 *Spatial distribution and clustering of bednet use in baseline population*

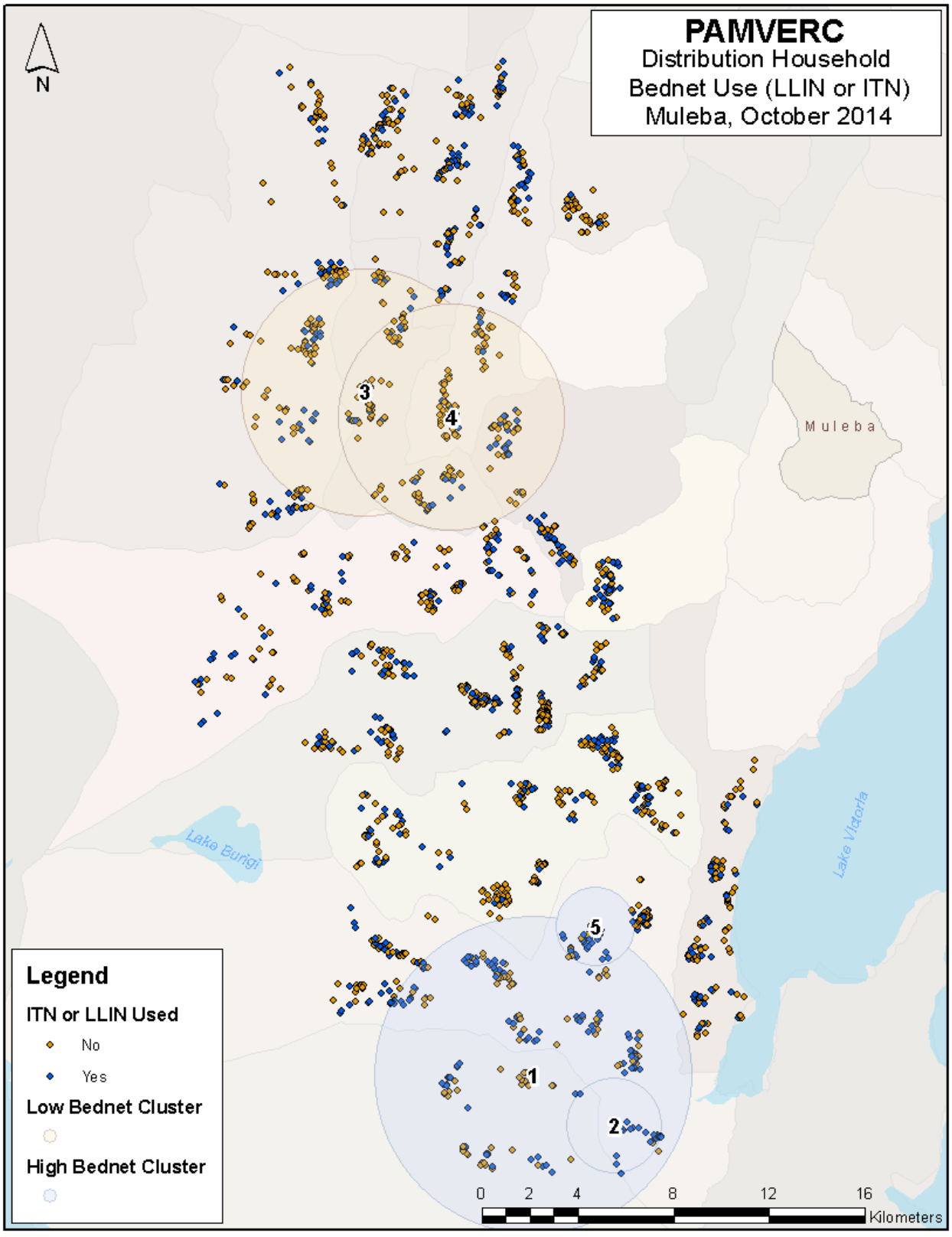


Figure 4.10 *Spatial distribution and clustering of indoor residual spraying in baseline population*

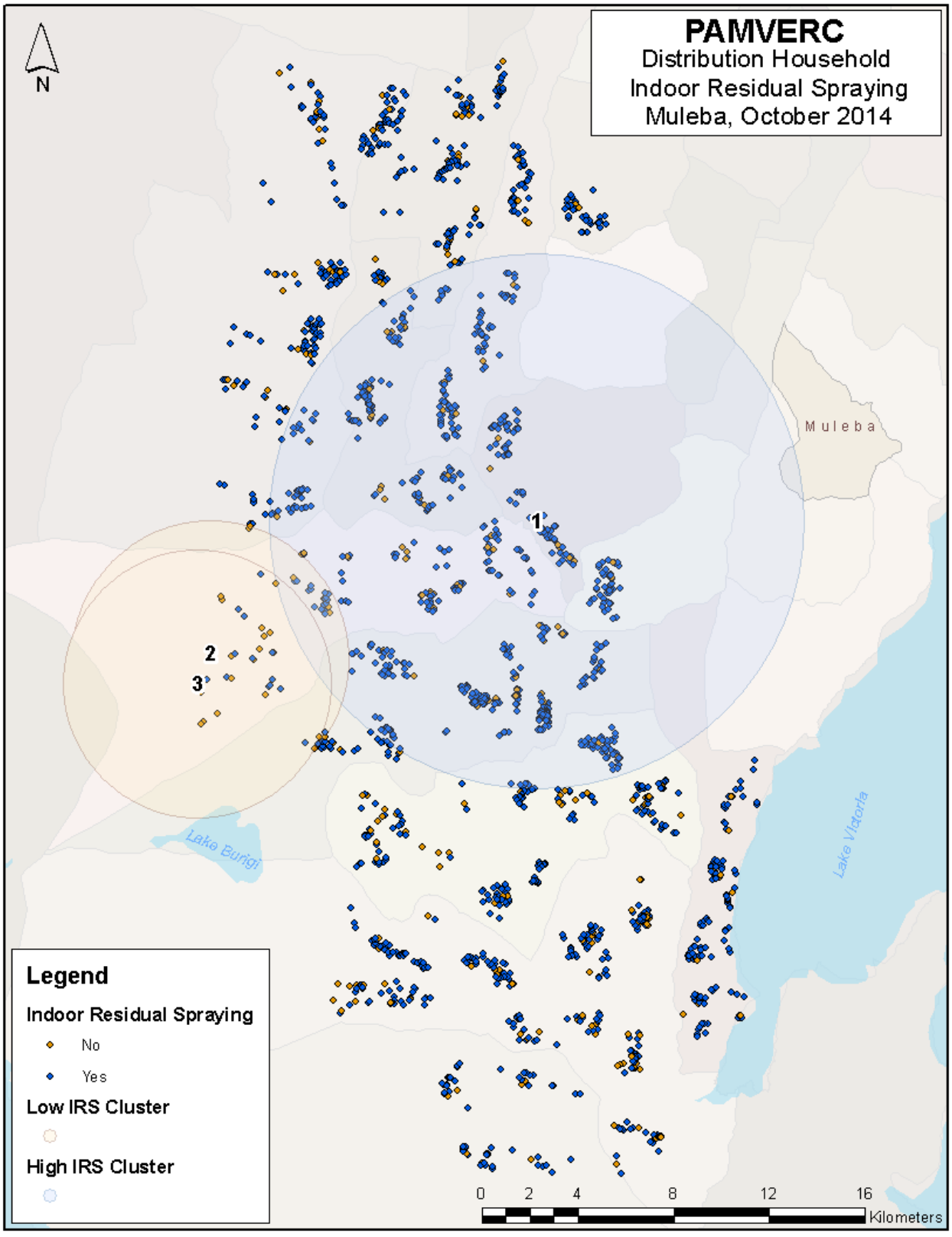


Table 4.3 *Baseline Odds of Malaria Infection (P. Falciparum or otherwise)*

	Simple Logistic		Adjusted Model	
Parameter	OR (95% CI)	p value	OR (95% CI)	p value
Age (years)	1.05 (1.03-1.06)	<0.0001	1.06 (1.04-1.08)	<0.0001
Elevation (per 100m)	0.57 (0.53-0.62)	<0.0001	0.56 (0.52-0.61)	<0.0001
Distance from Cluster Center (per 100m)	1.02 (1.01-1.03)	0.0036	1.01 (1.00-1.02)	0.1670
Household has Open Eaves	1.64 (1.43-1.89)	<0.0001	1.45 (1.24-1.70)	<0.0001
Roofing Material				
Grass/Leaves & Partial Metal	1.90 (1.57-2.30)	<0.0001	1.35 (1.07-1.72)	0.0603
Metal Sheets	REF		REF	
Head of Household Schooling				
No Education	2.50 (1.68-3.71)		1.78 (1.16-2.71)	
Primary Education	1.68 (1.15-2.46)	<0.0001	1.41 (0.95-2.11)	0.0062
Secondary/Tech or Higher Education	REF		REF	
Wealth Quintile				
Poorest	1.96 (1.57-2.45)		1.30 (1.01-1.68)	
Poorer	1.89 (1.52-2.34)		1.45 (1.14-1.83)	
Middle	1.31 (1.06-1.61)	<0.0001	1.06 (0.85-1.32)	0.0152
Less Poor	1.16 (0.94-1.42)		1.07 (0.86-1.33)	
Least Poor	REF		REF	
ITN/LLIN Used	0.82 (0.72-0.94)	0.0056	0.78 (0.67-0.90)	0.0007
Indoor Residual Spraying	0.92 (0.75-1.11)	0.3720		

Table 4.4 *Geographically-Weighted Adjusted Odds Ratios of Factors Predicting Malaria Infection (P. Falciparum or otherwise)*

	Odds Ratios					
Variable	Min	Lower Quartile	Median	Upper Quartile	Max	Range
Age	0.84	1.03	1.08	1.15	1.34	0.50
Elevation	0.00	0.12	0.25	0.46	29.96	29.96
Distance to Cluster Center	0.75	0.95	0.99	1.05	1.68	0.92
Household has Open Eaves	0.13	0.39	0.65	1.00	3.08	2.95
Roofing Material	0.02	0.53	0.90	1.54	6.86	6.84
Head of Household Schooling	0.06	0.61	0.87	1.36	4.71	4.65
Wealth Quintile	0.31	0.72	0.87	1.05	2.34	2.03
ITN/LLIN Used	0.05	0.48	0.74	1.25	26.22	26.17

Table 4.5 *Significant Clusters of Odds Ratios and Measures of Spatial Autocorrelation by Moran's I at Baseline*

Parameter	Cluster	Observed Mean	Expected Mean	O/E	p value	Moran's I Fit	Moran's I p value
Age						0.9858	<0.0001
High Odds Ratio Clusters	2	1.30	1.08	1.20	0.001		
	3	1.24	1.08	1.15	0.001		
	4	1.19	1.08	1.10	0.001		
Low Odds Ratio Clusters	1	1.03	1.12	0.92	0.001		
	5	0.95	1.09	0.87	0.001		
Elevation						0.9210	<0.0001
High Odds Ratio Cluster	1	19.23	0.41	46.90	0.001		
Distance from Cluster Center						0.9531	<0.0001
High Odds Ratio Clusters	1	1.38	0.99	1.39	0.001		
	5	1.12	1.00	1.12	0.003		
	4	1.08	1.00	1.08	0.002		
Low Odds Ratio Clusters	3	0.92	1.01	0.91	0.001		
	2	0.89	1.02	0.87	0.001		
Household has Open Eaves						0.9661	<0.0001
High Odds Ratio Clusters	1	2.02	0.70	2.89	0.001		
	3	2.18	0.79	2.76	0.001		
	4	2.04	0.80	2.55	0.001		
	9	1.40	0.81	1.73	0.001		
Low Odds Ratio Clusters	10	0.51	0.83	0.61	0.008		
	6	0.44	0.84	0.52	0.001		
	8	0.38	0.83	0.46	0.001		
	2	0.44	0.97	0.45	0.001		
	7	0.32	0.83	0.39	0.001		
	5	0.27	0.83	0.33	0.001		
Roof						0.9465	<0.0001
High Odds Ratio Clusters	2	4.94	1.16	4.26	0.001		
	1	4.33	1.13	3.83	0.001		
	5	4.03	1.18	3.42	0.001		
	3	2.78	1.14	2.44	0.001		
	7	1.86	1.17	1.59	0.001		
Low Odds Ratio Clusters	6	0.77	1.30	0.59	0.001		
	4	0.75	1.36	0.55	0.001		
Head of Household Schooling						0.9573	<0.0001
High Odds Ratio Clusters	1	3.06	1.03	2.97	0.001		
	2	2.78	1.06	2.62	0.001		
	4	2.03	1.08	1.88	0.001		
	6	1.77	1.09	1.62	0.001		
Low Odds Ratio Clusters	5	0.83	1.20	0.69	0.001		
	3	0.55	1.22	0.45	0.001		
Wealth Quintile						0.9655	<0.0001
High Odds Ratio Cluster	1	1.50	0.86	1.74	0.001		
	3	1.19	0.88	1.35	0.001		
Low Odds Ratio Clusters	2	0.78	1.01	0.77	0.001		
ITN/LLIN Used						0.5810	<0.0001
High Odds Ratio Cluster	1	13.16	0.97	13.57	0.001		
	2	3.73	0.97	3.85	0.011		

Figure 4.11 *Geographically-weighted adjusted odds ratios of age associated malaria infection at baseline*

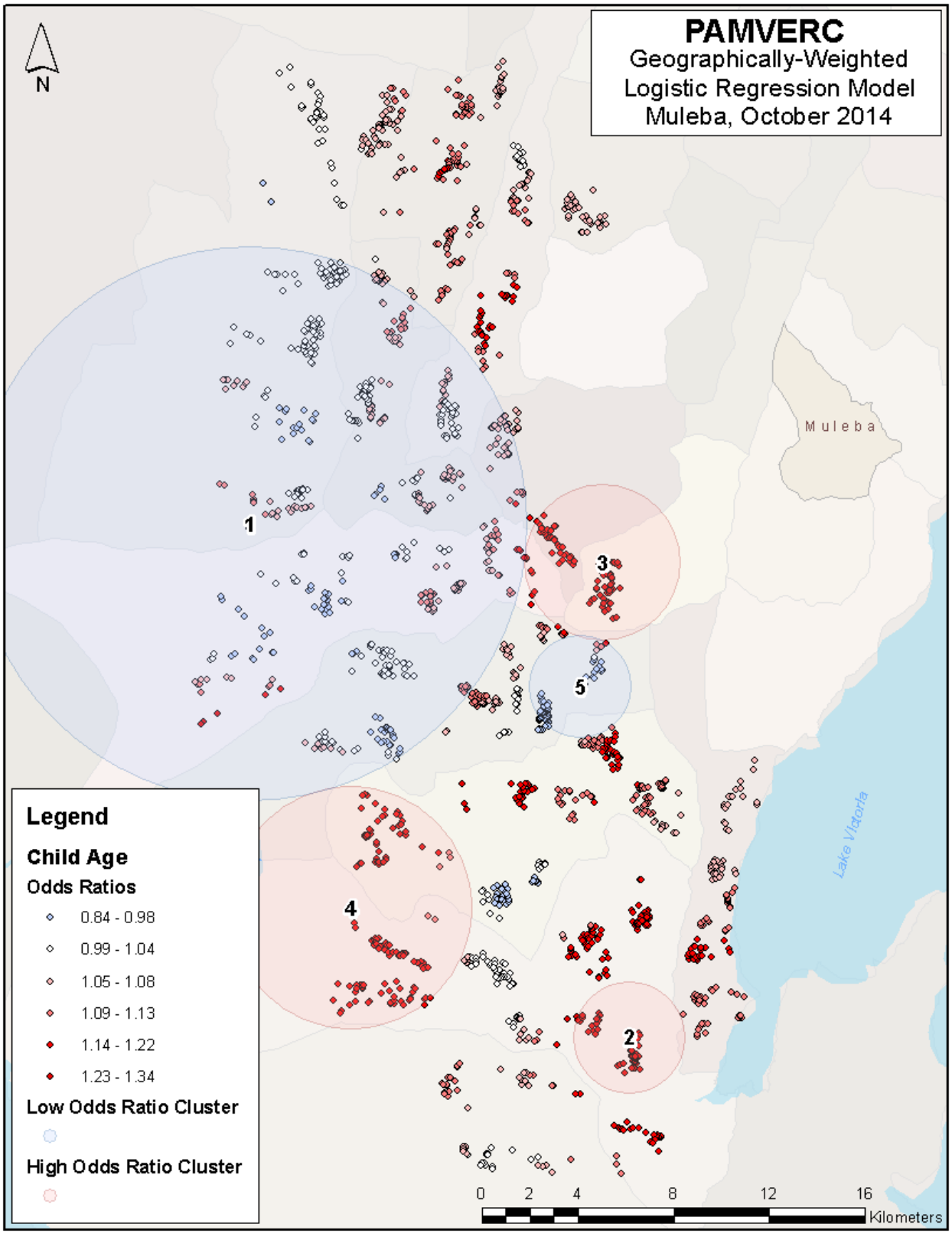


Figure 4.12 *Geographically-weighted adjusted odds ratios of elevation associated malaria infection at baseline*

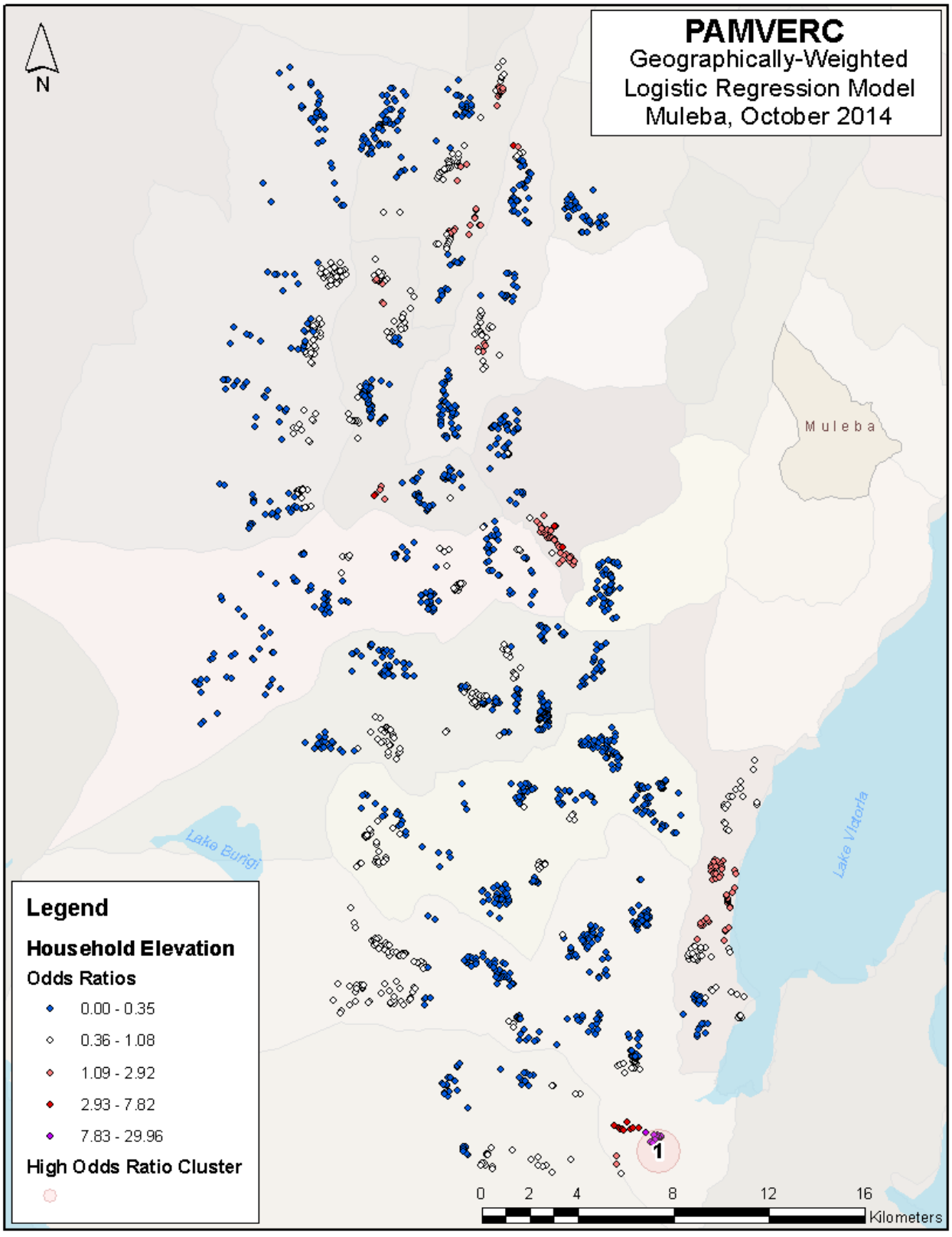


Figure 4.13 *Geographically-weighted adjusted odds ratios of distance to cluster median center associated malaria infection at baseline*

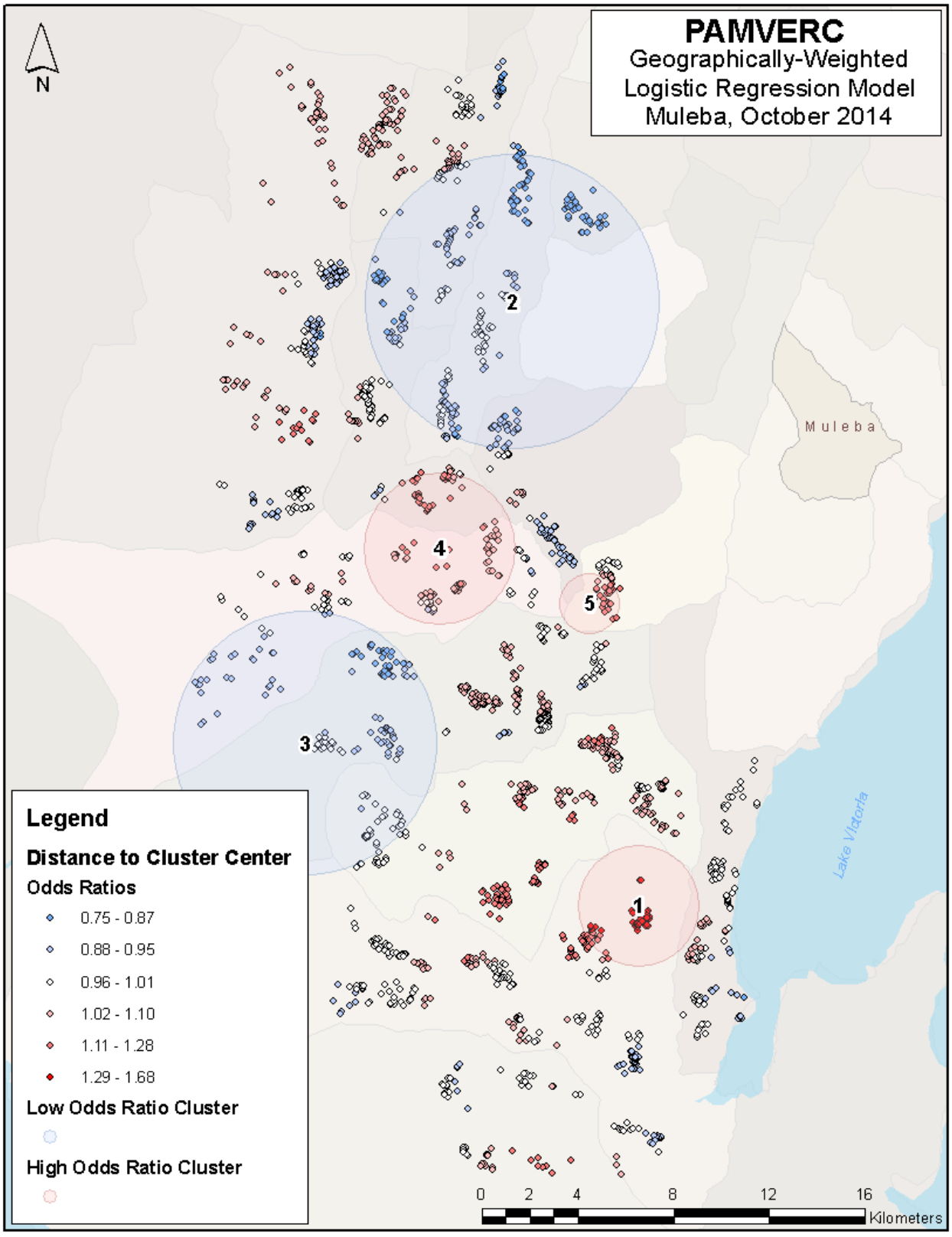


Figure 4.14 *Geographically-weighted adjusted odds ratios of open eaves associated malaria infection at baseline*

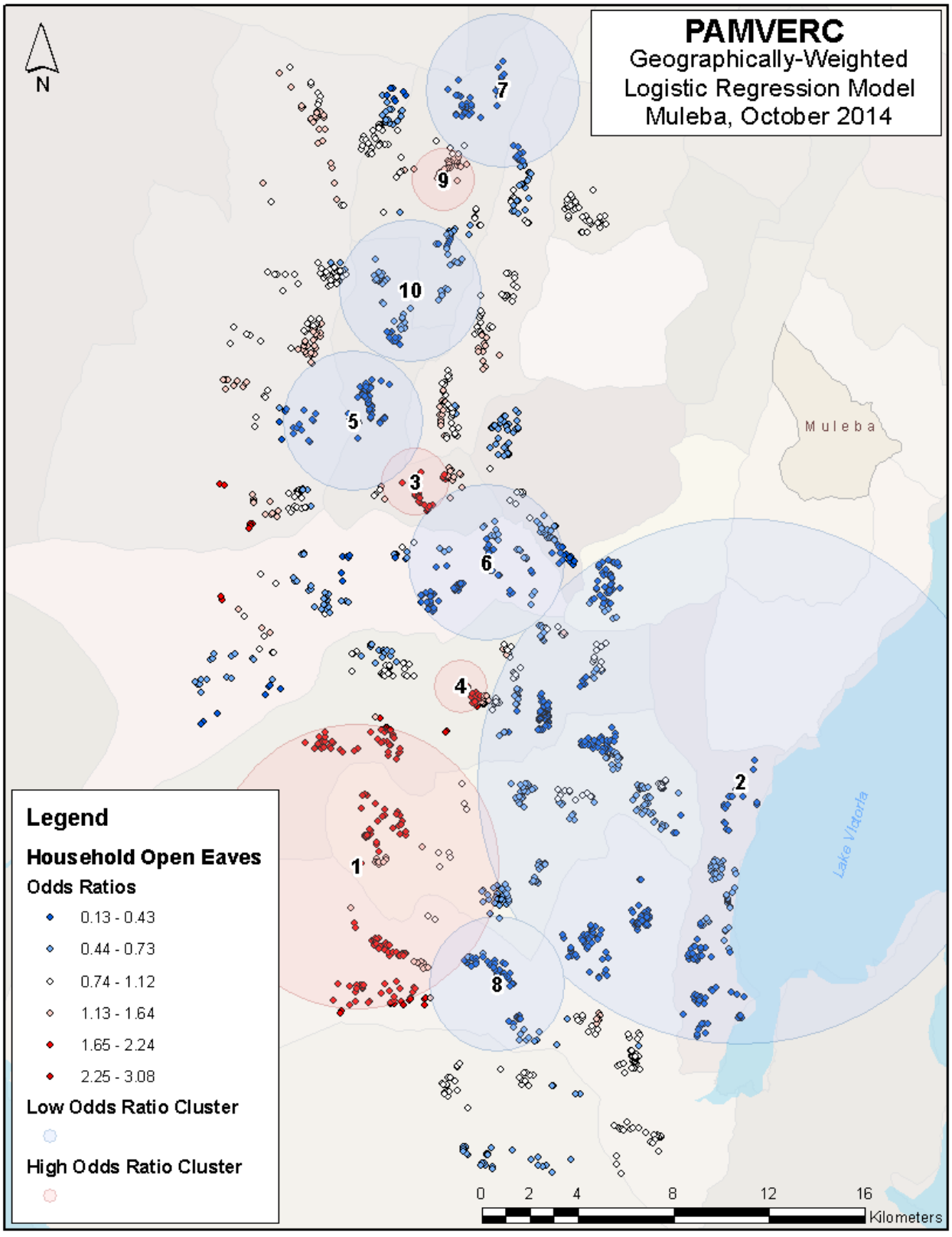


Figure 4.15 *Geographically-weighted adjusted odds ratios of roofing material associated malaria infection at baseline*

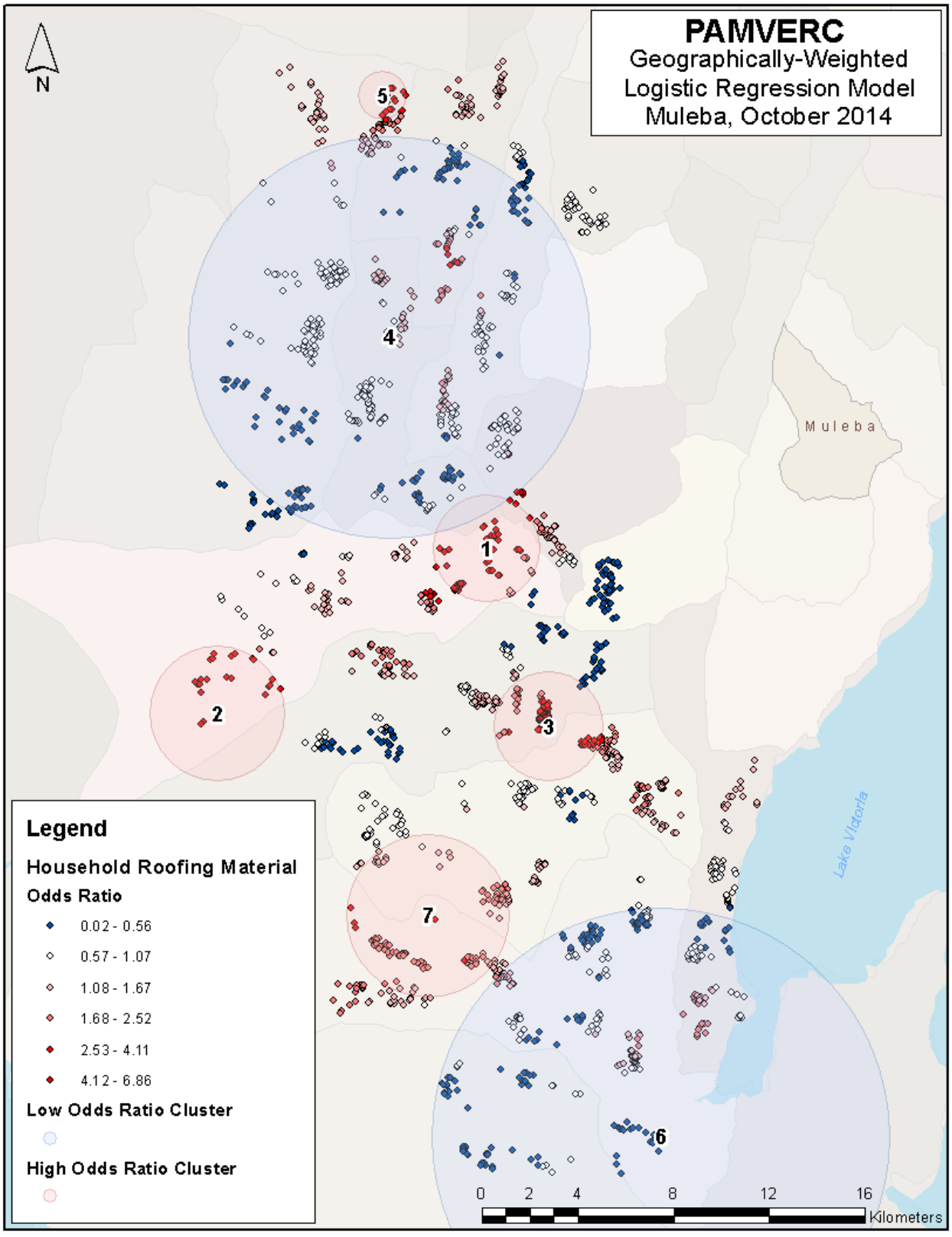


Figure 4.16 *Geographically-weighted adjusted odds ratios of head of household schooling associated malaria infection at baseline*

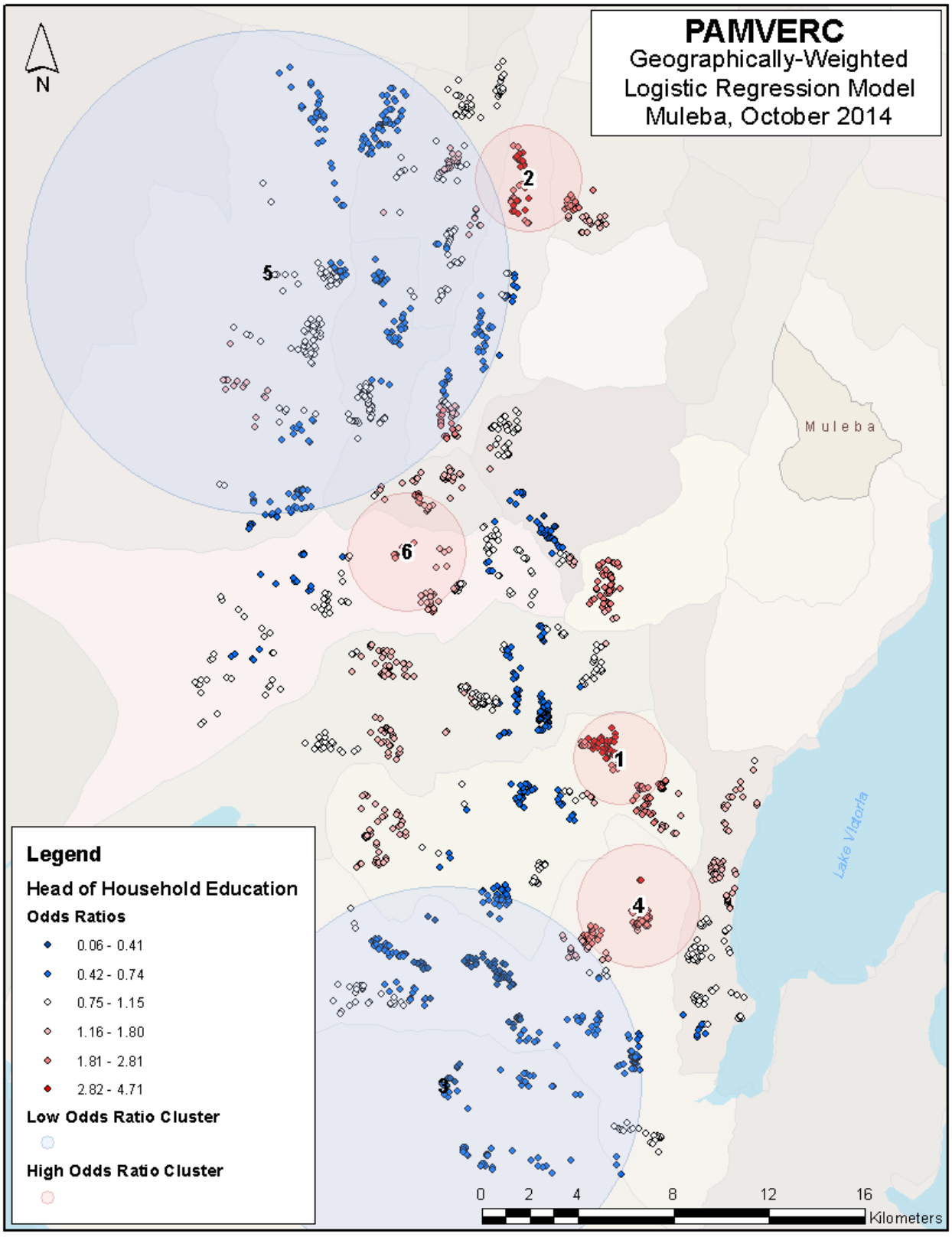


Figure 4.17 *Geographically-weighted adjusted odds ratios of socio-economic status associated malaria infection at baseline*

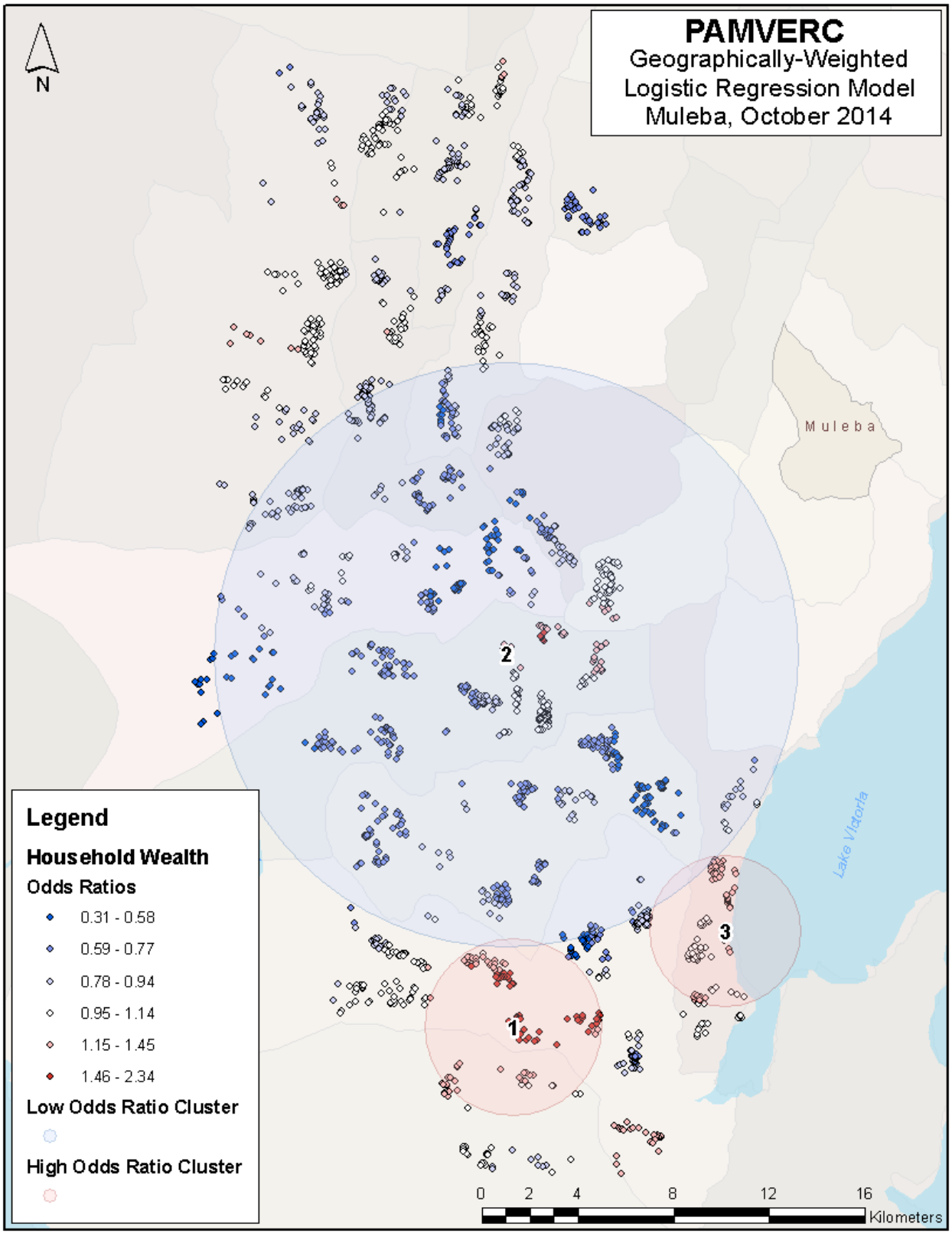
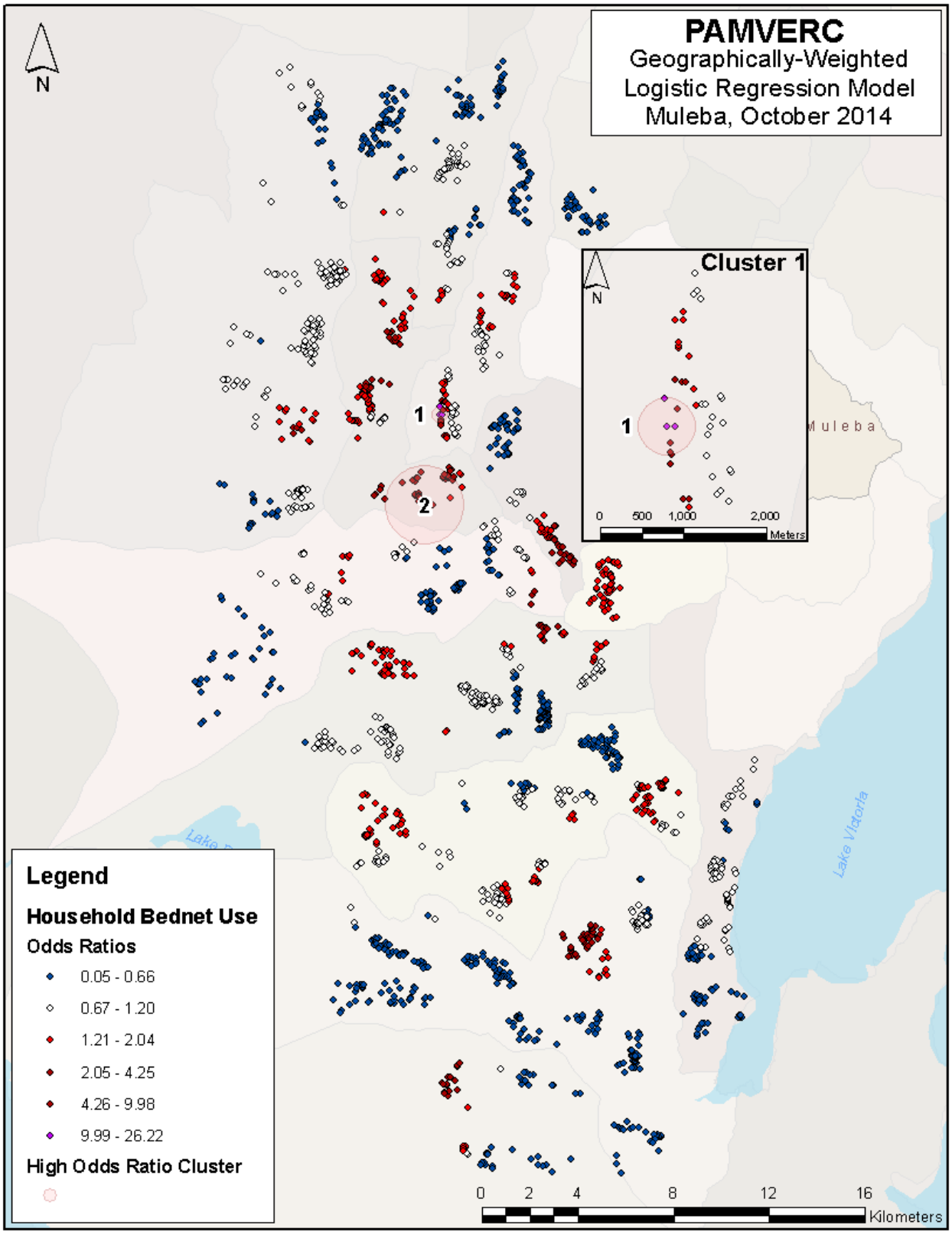


Figure 4.18 *Geographically-weighted adjusted odds ratios of bednet use associated malaria infection at baseline*



5. Chapter 2 – Inter-Cluster Contamination

A Semivariance Analysis of Community Effect Ranges of Malaria Vector Control Interventions in a Four-Armed Malaria Trial in Muleba, Tanzania

Charles Russell Thickstun, Natacha Protopopoff, Ann Jolly, Alice Zwerling, Manisha Kulkarni

Formatted in preparation for submission to

The Malaria Journal

5.1 Abstract

Background The presence of community effects in cluster randomized trials of malaria vector control interventions has led to the implementation of "buffer zones" around clusters to limit the potential for contamination between interventions. No consensus has been reached on how large these buffers need to be in order to encapsulate the effect.

Objectives This study aims to determine the presence and range of a community effect on household-level malaria infection in trial clusters four months post-intervention and to compare the calculated range to the current 300m buffer rule. It further aims to identify differences in the range of community effect between intervention arms with and without Indoor Residual Spraying.

Methods Geographically-weighted adjusted odds of malaria infection in children aged 6 months to 14 years were determined four months following the implementation of a novel malaria intervention in Muleba, Tanzania. Robust semivariograms were calculated for each of 48 intervention clusters and fit to semivariance models by Weighted Least Squares.

Results 6,440 children from 2,785 households were included in the GWLR. Prevalence of *Plasmodium falciparum* infection was 45.9% in the study population. Twenty (20) clusters had significant residual effect ranges and Sine Hole Effect models were fit for 13 of those indicating periodicity in the study area. Effective range estimates for the whole study area had a median value of 1,210 meters (IQR: 958 – 1,691). Clusters with IRS had a higher median range value: 1,535m (IQR: 976 – 3,398) than those without IRS: 1,168m (IQR: 829 – 1,504).

Conclusions A community effect was present in 20/48 clusters as determined by WLS. The community effect of interventions extended a median average of 600m beyond current buffer areas for all clusters, and 800m for IRS interventions. Future studies should aim to maximize buffer distances between intervention clusters to minimize contamination between intervention arms.

Keywords Malaria, Community Effect, Buffer, Semivariance, Spatial Analysis

5.2 Background

Cluster randomized controlled trials (cRCTs) are frequently used to evaluate the effectiveness of Long-Lasting Insecticidal Net (LLINs) and Indoor Residual Spraying (IRS) interventions on clinical and entomological outcomes in many different malaria transmission settings. In these trials, villages are randomly allocated into intervention groups rather than individuals (Friedman, Furberg, & DeMets, 2010). This allows for easier distribution of the trial interventions and creates a better simulation of a real-world rollout of a large scale malaria interventions for estimation of effects.

This clustering of interventions results in an additional effect through localized reductions in malaria transmission, and previous studies have noted the presence of a community effect to vector control cRCTs (Lines, Myamba, & Curtis, 1987; Maxwell, Myamba, Njunwa, Greenwood, & Curtis, 1999; Maxwell, et al., 2002; Gimnig, et al., 2003; Howard, et al., 2000). The community effect extends the benefits of high community-level intervention coverage to protect individuals in the community who do not themselves use LLINs or IRS. Spatially, it may extend the intervention effects outside of the community boundaries, resulting in partial reduction of malaria transmission in neighbouring communities or households. This amplification or extension of the intervention effects beyond the intervention areas suggests that the interventions are more impactful than initially believed, as they provide protective effects to larger areas than they are deployed to. However, this community effect comes with a drawback when evaluating interventions in a cRCT trial, as the spillover effect between intervention groups can cause between-group differences in intervention effects to trend towards the null hypothesis, reducing the significance of and potentially discarding effective interventions. The WHO recommends the use of a buffer zone or "fried egg" approach, where samples are only drawn from the center of cluster areas in order to minimize this possibility (VCAG, 2017). The size and composition of these buffer zones, however, is not yet included in any guidance provided. This has led to studies instituting buffer rules ranging anywhere from 300 meters (Hawley, et al., 2003) to 1500 meters (Bousema, et al., 2016). Establishing a consistent measure of the effect distance in malaria vector-control cRCTs is necessary to determine how large these buffer zones should be in future trials.

In order to obtain a continuous measure for the range of effect in vector control trials this paper constructs a fitted semivariogram model for each trial intervention cluster using the

residual odds of malaria from a geographically-weighted logistic regression model. This seeks to capture the intervention effect in the residual odds of malaria, and describe the distance within which the variation in intervention effects between households increase as a function of distance, while controlling for likely confounders, and establish an effective range where the local effect ceases.

5.3 Methods

5.3.1. Data Source

This study uses data from the first post-intervention cross-sectional survey conducted four months after intervention distribution in 2015 as part of a cluster-randomized trial in Muleba, Tanzania. This trial was led by the Pan-African Malaria Vector Research Consortium (PAMVERC), a collaboration between the London School of Hygiene & Tropical Medicine (UK), Kilimanjaro Christian Medical University College (Tanzania) and the National Institute for Medical Research (Tanzania). The four-armed intervention study investigates the efficacy of a conventional Olyset LLIN, an Olyset Plus LLIN, a conventional Olyset LLIN in combination with a novel IRS compound, and an Olyset Plus LLIN in combination with a novel IRS compound. The conventional Olyset LLIN contains the current WHO recommended single pyrethroid (permethrin), while the Olyset Plus LLIN includes permethrin plus piperonyl butoxide (PBO), a synergist which slows the oxidative metabolic process in mosquitoes by stopping enzymes that would otherwise break down the permethrin (NPIC, 2017). The new IRS formulation is Actellic®300CS (micro-encapsulated pirimiphos-methyl), a non-permethrin-based insecticide (WHO, 2013). Complete trial methods and results are discussed elsewhere (Protopopoff, et al., 2018).

Data were collected from a random sample of households in each cluster using interviewer-guided questionnaires with questions on demographic, and socioeconomic status indicators. The data includes a follow-up clinical and parasitological component which includes Rapid Diagnostic Testing (RDT) results for malaria, temperature, and hemoglobin levels, for randomly selected children within the selected households.

5.3.2. Study Area

All data collected in this trial is geocoded in World Geodetic System 1984 (WGS_1984) to allow for geospatial analyses. Initial mapping survey coordinates for each household were imported to ESRI ArcGIS 10.4 (ArcGIS, 2015) and projected into Africa Equidistant Conic projected coordinate system, preserving distances between points as accurately as possible for spatial analyses. The spatial extent of the study covers 1433km^2 between 657,023m and 686,018m east of meridian 25 and 187,255m to 236,686m south of the equator and includes 29,311 surveyed households in the Muleba district of the Kagera region of Tanzania. Elevation of study households ranges from 1075m to 1654m above sea level. A series of hills bisects the study region with low-lying households occupying the western areas and along the shore of Lake Victoria in the southeast.

5.3.3. Study Population

The first post-intervention household survey was conducted July 2015 (4 months post intervention) and recorded 3,316 households. 3,282 of these matched to households in the initial mapping survey, permitting geocoding of the data. A total of 8,019 children completed a following clinical visit, of which 7,635 were able to be matched back to post-intervention

household survey data for geocoding. One thousand, one hundred and ninety-five (1,195) children were missing data on response or explanatory variables used in the principal component or logistic regression analyses, with 1,020 of those missing data on their RDT status. The final analyzed dataset comprised 6,440 children from 2,785 households across 48 intervention clusters.

5.3.4. Statistical Analyses

Descriptive statistics were developed for variables of interest in the post-intervention clinical cross-sectional sample of children. Age, sex, household elevation, distance to cluster center, presence of open eaves, household roofing material, head of household schooling, and wealth quintile were all expected to be likely predictors of malaria infection in children due to a review of the extant literature, and findings from a previous analyses of the baseline data from the same trial (Thickstun, Unpublished).[13] Univariable associations with malaria infection, defined as children that tested positive for malaria *Plasmodium spp.* parasites via rapid diagnostic test, were determined through logistic regression in SAS 9.4 (SAS, 2012). Variables with significant association ($p<0.05$) were selected for a preliminary multivariable logistic regression model establishing the odds of malaria infection excluding variables which were inherently spatial in nature as they are likely to capture most of the effect we are hoping to observe in the residual values. All selected variables were checked for spatial autocorrelation through the calculation of a Moran's I statistic in SAS 9.4. The software GWR4 (GWR4, 2015) was used to construct a multivariable geographically-weighted logistic regression model. Residual values were extracted from the results of the GWLR for each datapoint and the dataset

[13] Sex was not included in the baseline clinical questionnaire and was therefore unable to be included previously.

was subdivided into clusters as assigned for intervention allocation. This division of the dataset into clusters is beneficial twofold: 1) it allows for the creation of quasi-stationarity in the cluster area as a sub-division of the greater study (Journel & Huijbregts, 1978), and 2) it allows for the assessment of cluster means individually, minimizing any spatial effect that could occur between clusters irrespective of cluster intervention.

Lag distances were calculated for each cluster to ensure a minimum of 50 data pairs per lag distance (Chilès & Delfiner, 2012) and pairs whose paired distance exceeded one half the total bounded data distance in each cluster were excluded from this analysis (Journel & Huijbregts, 1978). A Moran's I was calculated in SAS 9.4 to assess the spatial autocorrelation for the residual odds of malaria for each cluster dataset and the robust semivariance was calculated in SAS 9.4. A one-way analysis of variance (ANOVA) was conducted in SAS 9.4 to establish if there were any significant differences in the analysis characteristics for each allocation group.

5.3.5. Semivariance Analysis

Tobler's first law of geography states that correlation is a function of distance, with near samples tending to be more similar than distant ones (Tobler, 1970). A semivariogram is a model of this spatial autocorrelation in which the semivariance of measurements, or one half the mean squared difference in a set of measurements from two locations, are plotted against the distance separating those points (Cressie, 1993). This allows for a semivariogram curve to be fitted, originating at the origin and extending upwards as a function of distance.

Robust semivariance estimates are then fit to Exponential, Gaussian, Spherical, and Sine Hole Effect model forms using SAS 9.4. In cases where all 4 model structures were deemed to be a questionable fit, a Power model was considered. Significant range parameter estimates from

the model of best-fit were used to calculate an "Effective Range" for each cluster, according to the parameter estimate of the corresponding model.[14] These resulting Effective Ranges were grouped by intervention allocation and compared to the 300 meter spillover buffer currently being used.

5.4 Results

5.4.1. Characteristics of Study Population

Table 5.1 describes the characteristics of the study population. A total of 2,953 children tested positive for malaria parasites (*P. falciparum* or other) for an overall prevalence of 45.9%. Children who tested positive for malaria parasites were older on average (7.60 years, 95%CI:7.46 – 7.73) than those who tested negative (6.17 years, 6.04 – 6.03), lived at lower elevations (1,298 meters, 1,295 – 1,301 vs 1,335 meters, 1,331 – 1,338), and further away from cluster centers (1,430m, 1,397 – 1,462 vs 1,234m, 1,207 – 1,262). Malaria infection showed a moderate association with natural roofing materials (55.63% vs 44.62%), and very strong associations with male sex (47.21% vs 44.56%), households with open eaves (50.34% vs 37.87%), low head of household education level (51.31% vs 27.27%), and low household wealth (54.90% vs 38.14%). All variables of interest had a significant spatial trend in the study population (Moran's I $p<0.0001$) **(*Table 5.2*)**.

[14] Gaussian and Exponential functions both approach the sill asymptotically and therefore do not have a "true" range estimate. As such, SAS 9.4 considers the Effective Range to be the point at which the function reaches 95% of the sill (SAS, 2014).

5.4.2. Logistic Regressions

Table 5.3 displays the crude and adjusted logistic regression results for each variable of interest. In the adjusted model increased age (OR = 1.12, 1.10 – 1.13), being male (OR = 1.15, 1.04 – 1.28), living in a household with open eaves (OR = 1.51, 1.35 – 1.69), in a house with natural roofing materials (OR = 1.17, 1.01 – 1.35), lower head of household education (ORs = 1.83, 1.32 – 2.53; 1.63, 1.19 –2.23), and in lower wealth quintiles (ORs = 1.46, 1.21 – 1.74; 1.27, 1.07 – 1.50; 1.20, 1.02 – 1.41; 1.16, 0.98 – 1.36) were all associated with higher odds of malaria infection. Living at higher elevation (OR = 0.65, 0.62 – 0.69 per 100m) was associated with lower odds of malaria infection.

The adjusted model described above was used to compute a GWLR model. At 439 nearest-neighbours the GWLR provided a better goodness-of-fit when compared to the global logistic model (AICc 7,187 vs 8,239). Local odds ratio estimates for each variable can be seen in ***Table 5.4*** with eave type, household roofing material, and head of household schooling showing the greatest variability.

5.4.3. Cluster Characteristics

Table 5.5 lists the analysis characteristics of each trial cluster in the study area, as well as any residual spatial clustering that may exist in the GWLR residuals. There were no significant differences in cluster population (p=0.9192), geographical size (p=0.0882), lag distances (p=0.6643), lag numbers (p=0.9405) or residual spatial clustering (p=0.9532) between intervention groups **(*Table 5.6*)**.

5.4.4. Semivariance Analysis

Fitted semivariance models for twenty (20) clusters had significant estimates for the residual odds of malaria effect range indicating a defined spatial trend. The remaining clusters did not have significant range estimates: two (2) clusters had fitted models with non-significant range estimates, twenty-two (22) fitted models had "Pure Nugget" effects suggesting no spatial trend, and four were fit with Power models and thus had a spatial trend that exceeded one half the maximum data range for estimation of the variogram. Of the clusters that were fit with significant range estimates, 5 were from LLIN clusters **(*Table 5.7*)**, 5 from LLIN_PLUS clusters **(*Table 5.8*)**, 5 from LLIN_PLUS+IRS clusters **(*Table 5.9*)**, and 5 from LLIN+IRS clusters **(*Table 5.10*)**.

Effective range estimates for the whole study area had a median value of 1,210 meters (IQR: 958 – 1,691). Clusters without IRS had a median value of 1,168m (IQR: 829 – 1,504), and clusters with IRS 1,535m (IQR: 976 – 3,398). Clusters with LLIN_PLUS and LLIN_PLUS+IRS had higher median effective range estimates than standard LLIN clusters at 1,190m (IQR: 829 – 1,538) versus 1,146m (IQR: 1,041 – 1,504) for non-IRS and 1,717m (IQR:940 – 2,815) versus 1,353m (IQR: 1,176 – 3,398) for IRS clusters **(*Table 5.11*)**.

5.5 Discussion

This study examined the distance of potential spillover, or contamination, effects between intervention clusters in a four-armed 48 cluster malaria vector control intervention by comparing the range of semivariance models across intervention groups. Results showed that nearly half of the clusters in the study area exhibited a clear spatial trend in the residual geographically-weighted odds of malaria infection. These results are agreement with previous studies that

suggest a community effect may be present in malaria vector control trials (Binka, et al., 1998; Howard, et al., 2000; Hawley, et al., 2003).

While previous studies have not been able to provide exact estimates with regard to the distance these community effects may extend to, this study has provided a detailed analysis of the effect ranges for each cluster in the study area (overall effect range 1,210 meters, IQR: 958 – 1,691). This suggests that the 300m buffer ranges used in this study (Protopopoff, et al., 2018) and in one malaria study in the western Kenya highlands (Hawley, et al., 2003) would be inadequate to prevent contamination between groups. One kilometer buffers, as was previously used in Muleba (West, et al., 2013) or 1,500 meters, as was employed in the western Kenyan highlands (Bousema, et al., 2016) would be much more appropriate in terms of encapsulating the majority of the effect. Note that the effective semivariance ranges are the absolute level of global variance[15], and as such the range may be reduced to "acceptable" levels of contamination for each study. Additionally, this analysis demonstrated that there is a difference between IRS (1,535m, IQR: 829 – 1,504) and non-IRS (1,168m, IQR: 976 – 3,398) intervention effect ranges, which suggests that more extensive buffer zones should be employed for cRCT vector control trials investigating the effects of indoor residual spraying.

The large discrepancies in effect range distance between clusters in the same intervention group is likely due to a combination of factors including the spatial arrangement of sampling points (and indeed cluster households) and unobserved factors that may be significantly associated with malaria and spatially trending that creates some unavoidable variability in the model. This study has used robust semivariance estimates to attempt to create a minimization of this effect (Cressie & Hawkins, 1980), as well as reporting median values and interquartile

[15] Or the asymptotic 95% sill in the case of Exponential or Gaussian models.

ranges to provide conservative estimates. While it is possible that there are indeed effects ranging as far as 4 kilometers from cluster households, these numbers should be considered in the picture of the group estimates as one might consider an outlying variable in any dataset (Hoaglin, Mosteller, & Tukey, 1983).

5.5.1. Nugget Effects and Power Models

This study excluded clusters with non-significant range estimates in the analysis of effective semivariance ranges. The inclusion of these estimates, as well as those fit with a "pure nugget" effect or Power model would cause the group semivariance range averages to skew in one of two directions: 1) clusters with no significant spatial trend or "pure nugget" effects could be considered as some form of suppression effect on the final estimated effective ranges, as these clusters suggest that the residual odds of malaria infection is not uniformly spatially trending, 2) clusters which were fit with a Power model could be considered as an amplification of the current effective range estimates, as these suggest that there is a substantial spatial trend in these clusters that extends past one half the bounded distance of the cluster size and is thus unable to be accurately estimated.[16] These two effects together create some form of adjustment of the effective range estimates, and methods of accounting for these effects should be the subject of future study.

5.5.2. Method of Determining Goodness-of-Fit

Residual cluster semivariance models were fit to the empirical semivariogram by weighted-least squares (Cressie, 1985). While this method is regularly proposed as an option for

[16] This is especially relevant as the Sine Hole Effect, which was fit most often to significant clusters approximates the power function at very small scales.

determining goodness-of-fit, the manual fitting option is usually suggested as the primary method (Journel & Huijbregts, 1978; Chilès & Delfiner, 2012) with some texts even leaving automated fitting as an afterthought (Banerjee, Carlin, & Gelfand, 2004). Because this paper was an exercise in hypothesis-testing it was thought that a standard automated fitting process would reduce the possibility of introducing any inherent bias from the author to the data, but there may still be information to be gleaned from further analysis of the empirical semivariance models in future research. All cluster semivariogram models with the WLS best-fit model drawn are included in the appendix.

A majority of clusters with significant range estimates were fit with Sine Hole Effect models **(*Table 5.12*)** suggesting that there may be some periodicity in the contamination effect. This is possibly due to the irregular sampling pattern which can cause random fluctuation in the data (Emery, 2007). SAS 9.4 attempts to correct for these fluctuations by adding a periodic (in this case sine) function to the semivariance model. Spatial variation should be mostly controlled for in the GWLR analysis, however this is not a perfect solution and spatial clustering of the residual odds of malaria still persists in the final dataset **(*Table 5.5*)**. Note that this spatial clustering aligns closely with the clusters for which no spatial trend was observed, although there are some notable exceptions **(*Figure 5.1*)**.

5.6 Conclusion

Data from this analysis suggests that there may be a substantially larger community effect to vector-control interventions than previously thought. The estimated ranges of intervention spillover effects on malaria infection from a trial in Muleba, Tanzania, which reflects the distance over which malaria vector control interventions have an impact on malaria prevalence,

extend substantially beyond current best practices for trial cluster buffer sizes, especially in intervention clusters utilizing IRS. This potential contamination of between group differences in interventions may be causing a trend towards the mean in significant difference estimates in cluster-randomized vector control trials, and therefore may be impacting the capacity of trials to deliver meaningful results on the interventions under study. Future research should be incorporated into malaria vector control trials to estimate range effects for different types of LLIN and IRS interventions, and account for these effects in the interpretation of trial results.

5.7 References

ArcGIS. (2015). ArcGIS 10.4.1. Redlands, CA, USA: Envicornmental Systems Research Institute (ESRI) Inc.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical Modelling and Analysis for Spatial Data.* New York: Chapman & Hall.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G. H., . . . Smith, P. G. (1998). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health, 1*(2), 147-154.

Bousema, T., Stresman, G., Baidjoe, A. Y., Bradley, J., Knight, P., Stone, W., . . . Cox, J. (2016). The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLOS Medicine, 13*(4).

Chilès, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty* (Second ed.). John Wiley & Sons Inc.

Cressie, N. (1985). Fitting Variogram Models by Weighted Least Squares. *Mathematical Geology, 17*(5), 563-586.

Cressie, N. (1993). *Statistics for Spatial Data.* New York: John Wiley & Sons, Inc.

Cressie, N., & Hawkins, D. M. (1980). Robust Estimation of the Variogram: I. *Mathematical Geography, 12*(2), 115-125.

Emery, X. (2007). Reducing fluctuations in the sample variogram. *Stcochastic Environmental Research and Risk Assessment, 21*, 391-403.

Friedman, L. M., Furberg, C. D., & DeMets, D. L. (2010). *Fundamentals of clinical trials* (4th ed.). St. Louis: Springer.

Gimnig, J. E., Kolczak, M. S., Hightower, A. W., Vulule, J. M., Schoute, E., Kamau, L., . . . Hawley, W. A. (2003). Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *The American Journal of Tropical Medicine and Hygiene, 68*(4 Suppl), 115-120.

GWR4. (2015, 03). GWR4 v4.0.9. Maynooth , Ireland: National Center for Geocomputation, National University of Ireland Maynooth and Department of Geography, Ritsumeikan University, Japan.

Hawley, W. A., Phillips-Howard, P. A., Ter Kule, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., . . . Hightower, A. W. (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *American Journal of Tropical Medicine and Hygiene, 68*(Suppl 4), 121-127.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis.* New York: Wiley.

Howard, S. C., Omumbo, J., Nevill, C., Some, E. S., Donnelly, C. A., & Snow, R. W. (2000). Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of The Royal Society of Tropical Medicine and Hygiene, 94*(4), 357-360.

Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics.* New York: Academic Press.

Lines, J. D., Myamba, J., & Curtis, C. F. (1987). Experimental hut trials of permethrin-impregnated mosquito nets and eave curtains against malaria vectors in Tanzania. *Medical and Veterinary Entomology, 1*, 37-51.

Maxwell, C. A., Msuya, E., Sudi, M., Njunwa, K. J., Carneiro, I. A., & Curtis, C. F. (2002). Effect of community-wide use of insecticide-treated nets for 3–4 years on malarial morbidity in Tanzania. *Tropical Medicine and International Health, 7*(12), 1003-1008.

Maxwell, C. A., Myamba, J., Njunwa, K. J., Greenwood, B. M., & Curtis, C. F. (1999). Comparison of bednets impregnated with different pyrethroidsfor their impact on mosquitoes and on re-infection with malaria after clearance of pre-existing infections with chlorproguanil-dapsone. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 93*, 4-11.

NPIC. (2017). *Piperonyl Butoxide General Fact Sheet.* Oregon State University, National Pesticide Information Center. Retrieved from npic.orst.edu/factsheets/pbogen.html

Protopopoff, N., Mosha, J. F., Lukaloe, E., Charlwood, J. D., Wright, A., Mwalimu, C. D., . . . Rowland, M. (2018). Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two fact. *Lancet, 391*, 1577-88.

SAS. (2012). SAS v9.4. Cary, NC, USA: SAS Institute Inc.

SAS. (2014). *SAS/STAT® 13.2 User's Guide.* Cary, NC, USA: SAS Institute Inc. Retrieved from SAS/STAT(R) 13.2 User's Guide.

Thickstun, C. R. (Unpublished). By All Odds: A Geographically-Weighted Logistic Regression Analysis of Risk Factors for Malaria in Muleba, Tanzania.

Tobler, W. R. (1970, June). A computer movie simulating urban growth in the detroit region. *Economic Geography, 46*(Suppl. Proceedings. International Geographical Union, Commission on Quantitative Methods), 234-240.

VCAG. (2017). *How to design vector control efficacy trials: Guidance on phase III vector control field trial design provided by the Vector Control Advisory Group.* Vector Control Advisory Group. Geneva: World Health Organization.

West, P. A., Protopopoff, N., Rowland, M., Cumming, E., Rand, A., Drakeley, C., . . . Kleinschmidt, I. (2013). Malaria Risk Factors in North West Tanzania: The Effect of Spraying, Nets and Wealth. *PLOS One, 8*(6), 1-12.

WHO. (2013). *Report of the sixteenth WHOPES working group meeting.* Geneva: World Health Organization.

Table 5.1 *Characteristics of Post-Intervention Clinical Population by Rapid Diagnostic Test Result*

Parameter	Negative RDT n = 3487 μ (95%CI) or n (%)	Positive RDT n = 2953 μ (95%CI) or n (%)	p-value*
Age (years)	6.17 (6.04-6.30)	7.60 (7.46-7.73)	<0.0001
Elevation (meters)	1335 (1331-1338)	1298 (1295-1301)	<0.0001
Distance from Cluster Center (meters)	1234 (1207-1262)	1430 (1397-1462)	<0.0001
Sex			
Male	1664 (52.79)	1488 (47.21)	0.0372
Female	1823 (55.44)	1465 (44.56)	
Household has Open Eaves			
Yes	2048 (49.66)	2076 (50.34)	<0.0001
No	1439 (62.13)	877 (37.87)	
Roofing Material			
Grass/Leaves & Partial Metal	532 (44.37)	667 (55.63)	<0.0001
Metal Sheets	2955 (56.38)	2286 (44.62)	
Head of Household Schooling			
No Education	875 (48.69)	922 (51.31)	
Primary Education	2444 (55.39)	1968 (44.61)	<0.0001
Secondary/Tech or Higher Education	168 (72.73)	63 (27.27)	
Wealth Quintile			
Poorest	515 (45.10)	627 (54.90)	
Poorer	617 (50.95)	594 (49.05)	
Middle	745 (54.06)	633 (45.94)	<0.0001
Less Poor	734 (56.77)	559 (43.23)	
Least Poor	876 (61.86)	540 (38.14)	

Table 5.2 *Spatial Autocorrelation of Post-Intervention Parameters by Moran's I*

Parameter	**Moran's I**	
	Fit	**p value**
RDT	0.129	<0.0001
Age	0.0146	<0.0001
Sex	0.00778	0.0236
Household Elevation	0.9576	<0.0001
Distance to Center	0.872	<0.0001
Household Eaves	0.164	<0.0001
Household Roofing	0.161	<0.0001
Head of Household Schooling	0.142	<0.0001
Wealth	0.0983	<0.0001

Table 5.3 *Post-Intervention Odds of Malaria Infection (P. Falciparum or otherwise)*

	Simple Logistic		Adjusted Model	
Parameter	OR (95% CI)	p value	OR (95% CI)	p value
Age (years)	1.10 (1.08-1.11)	<0.0001	1.12 (1.10-1.13)	<0.0001
Sex (Male vs Female)	1.11 (1.01-1.23)	0.0328	1.15 (1.04-1.28)	0.0072
Elevation (per 100m)	0.66 (0.62-0.70)	<0.0001	0.65 (0.62-0.69)	<0.0001
Distance from Cluster Center (per 100m)	1.03 (1.02-1.03)	<0.0001	-	-
Household has Open Eaves	1.66 (1.50-1.85)	<0.0001	1.51 (1.35-1.69)	<0.0001
Roofing Material				
Grass/Leaves & Partial Metal	1.62 (1.43-1.84)	<0.0001	1.17 (1.01-1.35)	0.0368
Metal Sheets	REF		REF	
Head of Household Schooling				
No Education	2.81 (2.07-3.80)		1.83 (1.32-2.53)	
Primary Education	2.15 (1.60-2.88)	<0.0001	1.63 (1.19-2.23)	0.0008
Secondary/Tech or Higher Education	REF		REF	
Wealth Quintile				
Poorest	1.98 (1.69-2.31)		1.46 (1.21-1.74)	
Poorer	1.56 (1.34-1.83)		1.27 (1.07-1.50)	
Middle	1.38 (1.19-1.60)	<0.0001	1.20 (1.02-1.41)	0.0019
Less Poor	1.24 (1.06-1.44)		1.16 (0.98-1.36)	
Least Poor	REF		REF	

Table 5.4 *Post-Intervention Geographically-Weighted Adjusted Odds Ratios of Factors Predicting Malaria Infection (P. Falciparum or otherwise)*

	Odds Ratios					
Variable	Min	Lower Quartile	Median	Upper Quartile	Max	Range
Age	0.96	1.09	1.13	1.18	1.33	0.37
Sex	0.32	0.69	0.87	1.12	1.63	1.32
Elevation	0.08	0.31	0.50	0.68	2.29	2.21
Household has Open Eaves	0.50	1.10	1.45	1.87	3.73	3.23
Roofing Material	0.24	0.67	0.93	1.29	3.67	3.43
Head of Household Schooling	0.28	0.61	0.83	1.10	4.83	4.55
Wealth Quintile	0.64	0.82	0.88	0.95	1.20	0.56

Table 5.5 *Post-Intervention Cluster Intervention Allocation and Spatial Autocorrelation of Residual Odds of Malaria Infection (P. Falciparum or otherwise)*

Cluster	Allocation	n	Size (m)	Effective Size	Lag Distance	Lags	Moran's I	p value
1	LLIN	143	4621	2310	45	51	0.1620	0.0834
2	LLIN+IRS	126	4728	2364	50	47	0.1930	0.0412
3	LLIN+IRS	137	7123	3562	60	59	0.1530	0.1040
4	LLIN_PLUS+IRS	158	5702	2851	35	81	0.1100	0.1945
5	LLIN+IRS	122	8433	4217	70	60	0.1330	0.1517
6	LLIN+IRS	145	5312	2656	50	53	0.1860	0.0272
7	LLIN_PLUS	146	4156	2078	45	46	0.1140	0.1846
8	LLIN_PLUS	147	5749	2874	60	47	0.2800	0.0019
9	LLIN	145	5399	2700	40	67	0.2020	0.0291
10	LLIN_PLUS	51	5902	2951	400	7	0.2620	0.0203
11	LLIN	114	6708	3354	80	41	0.2760	0.0038
12	LLIN	169	7423	3712	60	61	0.1680	0.0341
13	LLIN_PLUS+IRS	170	6541	3270	50	65	0.1720	0.0218
14	LLIN	173	4128	2064	35	58	0.2010	0.0136
15	LLIN	121	5715	2857	105	27	0.0060	0.8613
16	LLIN+IRS	120	6696	3348	85	39	-0.0852	0.3963
17	LLIN+IRS	111	4981	2491	85	29	0.1330	0.1372
18	LLIN	105	4990	2495	85	29	0.2540	0.0122
19	LLIN+IRS	120	8336	4168	90	46	0.2920	0.0018
20	LLIN_PLUS+IRS	117	6990	3495	105	33	0.2800	0.0033
21	LLIN	109	5125	2562	65	39	0.2700	0.0018
22	LLIN_PLUS+IRS	98	3411	1705	55	31	0.0314	0.6744
23	LLIN_PLUS+IRS	135	6977	3489	85	41	0.0427	0.5591
24	LLIN_PLUS+IRS	116	5504	2752	85	32	0.3160	0.0005
25	LLIN+IRS	144	4635	2317	55	42	0.2690	0.0033
26	LLIN_PLUS	127	3933	1967	85	23	0.3060	0.0009
27	LLIN+IRS	134	10748	5374	415	12	0.0627	0.1710
28	LLIN_PLUS+IRS	131	5662	2831	45	62	0.1890	0.0299
29	LLIN_PLUS	145	2624	1312	35	37	0.2790	0.0018
30	LLIN_PLUS+IRS	138	4935	2467	55	44	0.1990	0.0249
31	LLIN_PLUS+IRS	140	7032	3516	80	43	0.2580	0.0014
32	LLIN_PLUS+IRS	146	2307	1154	30	38	0.0428	0.5812
33	LLIN	135	6166	3083	100	30	0.1040	0.2000
34	LLIN_PLUS	145	5043	2521	45	56	-0.0825	0.3835
35	LLIN	175	6643	3321	50	66	0.0763	0.3192
36	LLIN_PLUS	122	7154	3577	130	27	0.4160	<.0001
37	LLIN_PLUS	151	6522	3261	65	50	0.3110	0.0001
38	LLIN_PLUS	144	5167	2584	40	64	0.1280	0.1288
39	LLIN_PLUS+IRS	139	7942	3971	95	41	0.0345	0.5928
40	LLIN_PLUS	137	4839	2419	45	53	0.1010	0.2458

41	LLIN+IRS	147	11056	5528	110	50	0.2670	0.0015
42	LLIN_PLUS	137	6180	3090	50	61	0.4000	<.0001
43	LLIN+IRS	177	4729	2365	30	78	0.1510	0.0617
44	LLIN	145	5297	2648	55	48	0.1120	0.1882
45	LLIN	122	2995	1498	45	33	0.2960	0.0009
46	LLIN_PLUS+IRS	109	4489	2244	95	23	0.1520	0.0861
47	LLIN_PLUS	130	3439	1719	45	38	-0.0548	0.6159
48	LLIN+IRS	122	4748	2374	60	39	0.3220	0.0007

Table 5.6 *Post-Intervention Cluster Characteristics*

	LLIN	**LLIN_PLUS**	**LLIN_PLUS + IRS**	**LLIN + IRS**	**p value***
Mean Cluster Population	138	132	133	134	0.9192
Mean Cluster Size (m)	5434	5059	5624	6794	0.0882
Mean Lag Distance (m)	64	87	68	97	0.6643
Mean Number of Lags	45	42	45	46	0.9405
Number of Clusters with Residual Spatial Clustering	7	7	6	6	0.9532

*p-values were conducted by ANOVA or chi-square test as appropriate.

Table 5.7 *Model Fit and Range Estimates for LLIN Clusters*

Cluster	Form	Nugget		Scale		Range (meters)				Effective Range (meters)
		Est.	p value	Est.	p value	Est.	SE	95% CI	p value	
1	SPH	0.1765	.	0.017790	.	1146	311.41	520-1772	0.0006	1146
9	SPH	0.1877	<.0001	0.000001	0.9999	1341	0.00	1341-1341	.	.
11	GAU	0.1724	.	0.013840	.	1643	1044.23	0-3755	0.1237	.
12	SPH	0.1630	<.0001	0.010880	0.09	1824	0.00	1824-1824	.	.
14	SPH	0.1136	<.0001	0.031030	<.0001	1017	0.00	1017-1017	.	.
15	POW	0.0184	<.0001							.
18	SHE	0.2025	.	0.019100	.	1229	211.33	795-1663	<.0001	1229
21	SHE	0.1318	.	0.074900	.	1504	109.99	1281-1727	<.0001	1504
33	SHE	0.1706	.	0.037500	.	1041	79.61	878-1204	<.0001	1041
35	GAU	0.1938	<.0001	0.003824	0.4476	1651	0.00	1651-1651	.	.
44	EXP	0.1559	<.0001	0.065180	<.0001	1320	0.00	1320-1302	.	.
45	SHE	0.2165	.	0.029220	.	741	94.64	548-934	<.0001	741

Table 5.8 *Model Fit and Range Estimates for LLIN_PLUS Clusters*

Cluster	Form	Nugget		Scale		Range (meters)				Effective Range (meters)
		Est.	p value	Est.	p value	Est.	SE	95% CI	p value	
7	EXP	0.0724	<.0001	0.157300	<.0001	513	74.71	362-663	<.0001	1538
8	EXP	0.1940	<.0001	0.000001	0.9999	1406	0.00	1406-1406	.	.
10	SPH	0.1814	0.0015	0.000001	1	1402	0.00	1402-1402	.	.
26	SHE	0.0100	.	0.014000	.	829	21.79	784-875	<.0001	829
29	SPH	0.0776	<.0001	0.000001	0.9998	649	0.00	649-649	.	.
34	EXP	0.1876	<.0001	0.000001	0.9999	1260	0.00	1260-1260	.	.
36	SHE	0.1612	<.0001	0.013570	0.0437	1745	0.00	1745-1745	.	.
37	SHE	0.1306	<.0001	0.047660	<.0001	1190	38.14	1113-1267	<.0001	1190
38	EXP	0.2010	<.0001	0.000001	0.9999	1276	0.00	1276-1276	.	.
40	SHE	0.1022	<.0001	0.073930	.	44	0.42	43-44	<.0001	44
42	EXP	0.2077	<.0001	0.000001	0.9999	1523	0.00	1523-1523	.	.
47	SHE	0.1291	.	0.119100	.	1665	317.55	1021-2309	<.0001	1665

Table 5.9 *Model Fit and Range Estimates for LLIN_PLUS + IRS Clusters*

Cluster	Form	Nugget		Scale		Range (meters)				Effective Range (meters)
		Est.	p value	Est.	p value	Est.	SE	95% CI	p value	
4	EXP	0.1472	.	0.089710	.	1419	384.68	653-2185	0.0004	4257
13	GAU	0.1518	.	0.014010	.	1625	451.84	723-2528	0.0006	2815
20	GAU	0.1799	<.0001	0.000001	0.9999	1734	0.00	1734-1734	.	.
22	EXP	0.1676	.	0.042700	.	852	432.33	0-1736	0.0583	.
23	SPH	0.1665	<.0001	0.004684	0.5854	1738	0.00	1738-1738	.	.
24	SHE	0.1075	<.0001	0.081260	<.0001	940	24.13	891-990	<.0001	940
28	SHE	0.1948	<.0001	0.019010	0.0024	1393	0.00	1393-1393	.	.
30	POW	0.0637	<.0001							.
31	SHE	0.2006	.	0.020650	.	1717	414.84	879-2555	0.0002	1717
32	POW	0.0602	<.0001							.
39	SPH	0.2013	<.0001	0.000001	0.9999	1948	0.00	1948-1948	.	.
46	SHE	0.1373	<.0001	0.074930	<.0001	709	28.82	649-768	<.0001	709

Table 5.10 *Model Fit and Range Estimates for LLIN + IRS Clusters*

Cluster	Form	Nugget		Scale		Range (meters)				Effective Range (meters)
		Est.	p value	Est.	p value	Est.	SE	95% CI	p value	
2	SPH	0.1535	.	0.050070	.	1176	244.07	684-1667	<.0001	1176
3	SPH	0.1045	.	0.091940	.	1353	126.99	1098-1607	<.0001	1353
5	SPH	0.1937	<.0001	0.011930	0.3246	2101	0.00	2101-2101	.	.
6	SHE	0.1703	.	0.039860	.	976	77.03	821-1131	<.0001	976
16	POW	0.0774	<.0001							.
17	SPH	0.1580	<.0001	0.000001	0.9999	1237	0.00	1237-1237	.	.
19	EXP	0.2244	<.0001	0.028060	0.0314	2069	0.00	2069-2069	.	.
25	EXP	0.2276	<.0001	0.000001	0.9999	1153	0.00	1153-1153	.	.
27	SHE	0.1731	.	0.051710	.	3398	579.46	2107-4689	0.0002	3398
41	SPH	0.1411	<.0001	0.010460	0.0913	2746	0.00	2746-2746	.	.
43	GAU	0.1153	.	0.159000	.	2158	386.30	1388-2927	<.0001	3738
48	SPH	0.1274	<.0001	0.000001	0.9999	1171	0.00	1171-1171	.	.

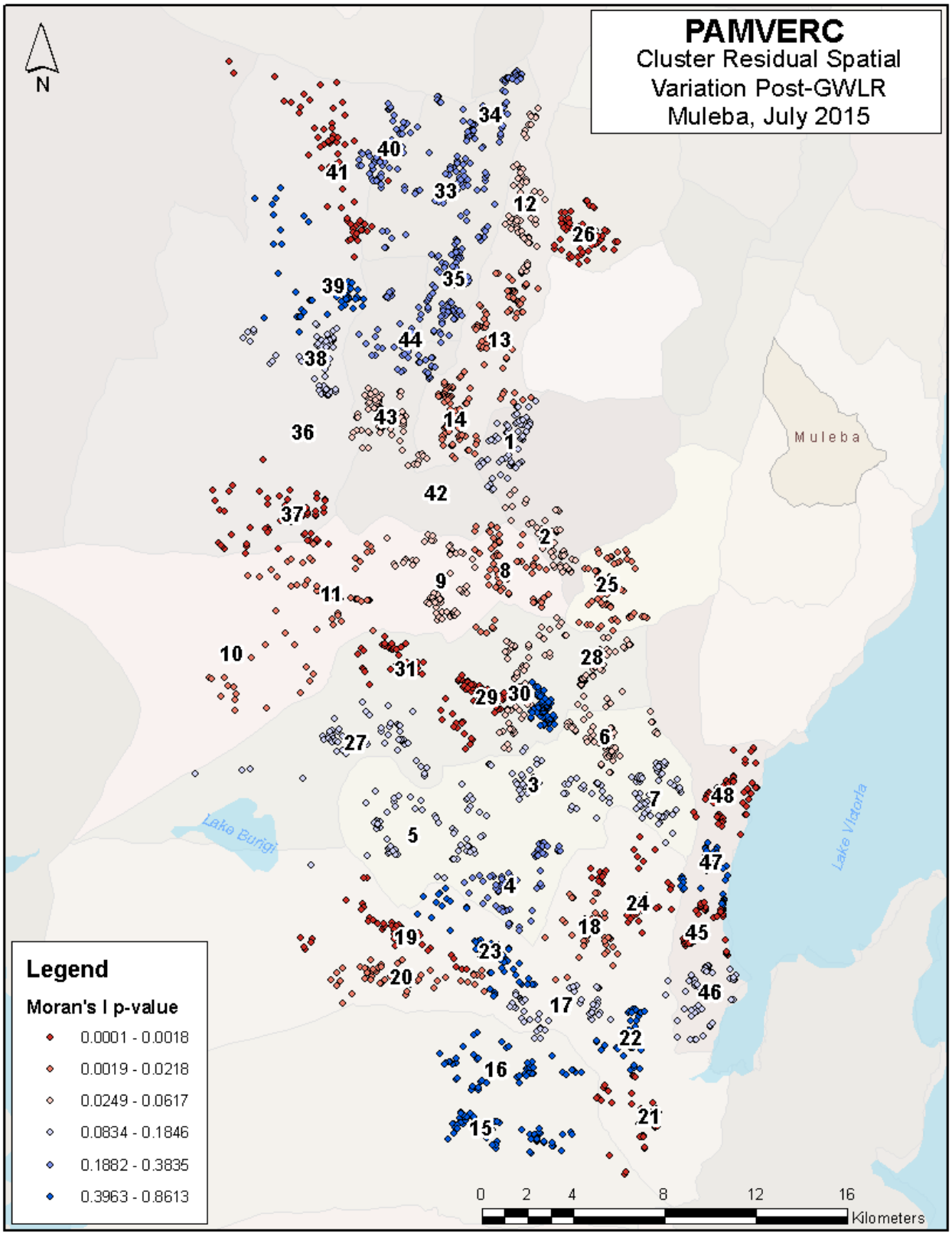
Table 5.11 *Effective Range Estimates of Residual Odds of Malaria Infection (P. Falciparum or otherwise) by Cluster Allocation*

	Effect Ranges (m)					
	Min	**Q1**	**Median**	**Q3**	**Max**	**Inter-Quartile Range**
Whole Dataset	**44**	**958**	**1,210**	**1,691**	**4,257**	**733**
Non-IRS	**44**	**829**	**1,168**	**1,504**	**1,665**	**675**
LLIN	741	1,041	1,146	1,229	1,504	188
LLIN_PLUS	44	829	1,190	1,538	1,665	291
IRS Clusters	**709**	**976**	**1,535**	**3,398**	**4.257**	**2,422**
LLIN_PLUS+IRS	709	940	1,717	2,815	4,257	1,875
LLIN+IRS	976	1,176	1,353	3,398	3,738	2,222

Table 5.12 - *Model Fit Type by Cluster Allocation*

	Significant Range Models				Non-Significant Range Models				
	SPH	**GAU**	**EXP**	**SHE**	**SPH**	**GAU**	**EXP**	**SHE**	**POW**
LLIN	1	0	0	4	3	2	1	0	1
LLIN_PLUS	0	0	1	4	2	0	4	1	0
LLIN_PLUS+IRS	0	1	1	3	2	1	1	1	2
LLIN + IRS	2	1	0	2	4	2	0	0	1
Total	**3**	**2**	**2**	**13**	**11**	**5**	**6**	**2**	**4**

Figure 5.1 *Cluster-specific Residual Spatial Clustering in Residual Geographically-Weighted Adjusted Odds of Malaria Infection*



6. Discussion

6.1 Results in Context

Tanzania remains a country at dramatically high-risk of malaria with 56,682,955 people deemed at "high risk" of infection (WHO, 2017). The spatial heterogeneity of malaria infection and the factors associated with it exhibited in this study add to an ever-increasing body of literature that has found similar effects. Even in just the last few years studies have established heterogeneity to malaria and its covariates across Africa in Senegal (Ndiath, et al., 2015), the Kenyan highlands (Bousema, et al., 2016; Homan, et al., 2016), Zimbabwe (Manyangadze, Chimbari, Macherera, & Mukaratirwa, 2017), the Gambia (Mwesigwa, et al., 2017), Malawi (Dear, et al., 2018), Uganda (Maziarz, et al., 2018), and Burkina Faso (Ouedraogo, et al., 2018). This explosion of spatial analysis has ranged from simple nearest-neighbour analyses (Binka, et al., 1996), to cluster detection (Brooker, et al., 2004), geographically-weighted regression (Haque, et al., 2012), discrete-time stochastic simulations (Silkey, et al., 2016), and likely more that fell outside the scope of this author's review of the literature.

While spatial methods have been used extensively in geoscience for decades following Noel Cressie's foundational work *Statistics for Spatial Data* (Cressie, 1960), spatial epidemiology has remained an underexplored field. This has created notable knowledge gaps in epidemiology, as highlighted by the hesitation by the WHO Vector Control Advisory Group (VCAG) to recommend a standard buffer size for cluster-randomized trials in their guidelines for phase III malaria vector control field trials (VCAG, 2017). The use of semivariance in this study to examine the residual odds of malaria infection post-intervention to provide a more continuous measure of spillover from clusters is, to the knowledge of this author, novel. Current trials

institute buffer zones anywhere from 300 – 1500 meters, greatly varying the cost of trials and leaving questions about the accuracy of results (Jarvis, Di Tanna, Lewis, Alexander, & Edmunds, 2017). While this thesis offers no good news on the monetary costs associated with trials, it does provide some semblance of direction going forward.

The primary focuses of this author in the near future will be to develop an anisotropic analysis to determine the directionality of the semivariogram models for individual clusters, in particular along gradients between clusters of different intervention allocations. Following this, the development of a raster surface of interpolated intervention effects will be done by two-dimensional Kriging. This will allow for the semivariance results to be interpolated across the study area, creating continuous geocoded estimates of contamination in the intervention area (Olea, 1974). Further establishing how intervention effects have spread across the study area, particularly in the context of emerging insecticide resistance, is vitally important in understanding the true effects of interventions in the study; failure to do so could result in the inappropriate rejection of trialed interventions.

6.2 Implications

This thesis suggests that there is enormous variability to factors associated with malaria infection, both in terms of passive risk factors and active interventions. The baseline (pre-intervention) geographically-weighted logistic regression analysis demonstrated that risk factors cannot be assumed as uniform across a study area, and that attempts must be made to further explore the interaction of all factors associate with malaria and their spatial locations. Going forward, studies should take care to consider the localized effects of various covariate factors in their analyses, as these may change across the study area. This may have implications not only

on the impact of controlling for variables in analysis, but also in the targeting of interventions for greatest effect.

Secondly this study has demonstrated that there are community effects of malaria vector control interventions that may extend past their cluster boundaries. These effects are fundamentally beneficial as they amplify the effect of the intervention beyond what is immediately suggested, but they also require that researchers take more care when evaluating between-group differences in intervention trials. These spillover or contamination effects will create a bias towards the null hypothesis in statistical analysis as they become inseparable as they bleed into each other. Until there is a preponderance of evidence as to the consistent size of these effects, vector control trials should attempt to maximize the buffer areas between studies to reduce the impact of this effect on the trial significance.

The replication of the geographically-weighted logistic regression in both the baseline and four months post-intervention datasets provides a point of comparison across the intervention period that did not have an applicable place in either study proper and has implications on the future of malaria studies. First, it can be noted that the variability in the geographically-weighted adjusted odds of malaria reduced dramatically with the introduction of the study interventions ***(Table 4.4 & Table 5.4)***. This is likely due to a combination of full coverage of the interventions – rather than the 59.80% with ITN/LLIN and 14.82% with IRS at baseline – with the introduction of novel insecticide products to combat resistance. In particular, in instances such as the baseline dataset where protective measures such as ITN/LLINs are shown to be risk-increasing should be carefully evaluated as they may be indicative of a failure of the intervention in the area that may not be externally relevant to other areas of the study.

Additionally, the presence of an increasingly significant effect from the Distance to Cluster Center variable on the adjusted odds of malaria in the post-intervention analysis – p-value Baseline = 0.1670 **(*Table 4.3*)** vs p-value Post-Intervention <0.0001 **(*Table 6.1*)** – further contributes to the suggestion that there are significant differences in malaria-associated variables in buffer areas as compared to core sampling areas. Note that in this study, as in many, the cluster boundaries are drawn by hand and therefore are not truly indicative of a rural-urban split **(*Table 6.1*, *Figure 6.1*)** and may therefore present difficulties in establishing between group differences in the core and buffer areas, especially with only a single cross-sectional examining households in the buffer. This thesis concurs with Silkey *et al*. (Silkey, et al., 2016) that vector control trials should endeavor to collect data on individuals uniformly across clusters, not just in core sampling areas, in order to provide more evidence from which to make estimations of these contamination effects and to validate those methods of doing so that are emerging. This author encourages other scholars to carefully examine the spatial aspects of their data when dealing with malaria – or indeed any vector-borne disease – as these aspects are both almost universally underexplored, and misunderstood. Only through the concerted effort to examine these aspects on a large scale can they be fully evaluated.

6.3 Conclusion

Despite the appeal of declaring that the risk factors for malaria are "set" – that there is no need to explore further mechanisms of action, rather than control for the same set of variables in every study – the introduction of spatial methods in the last decade has created a whole new perspective on the examination of malaria study areas. This thesis in particular has reinforced that there are enormous heterogeneities in every variable under study, and that interventions do

not apply quite as uniformly as they may be perceived to. While the practical realities of this effect seems to be skewed in our favor on the surface, as the presence of community effects indicate that large scale interventions will have greater malarial reduction impacts than they do when under study, these effects must be carefully evaluated to ensure that we are not improperly discarding interventions that have non-significant results simply due to a bias towards the null. The use of Geographically-Weighted Regression and introduction of Semivariance Analysis will hopefully provide useful tools to further our understanding of the disease and the factors associated with it.

While scholars and physicians have long bemoaned traditional beliefs about the causes of malaria as they present difficulties in the rollout of interventions (Desowitz, 1991), it should be carefully reflected upon that the mosquito-malaria theory was not experimentally confirmed until the end of the 19th century (Manson, 1900; Grassi, Bignami, & Bastianelli, 1899), and the common perception of malaria as quite literally "bad air" was not shaken until well after even amongst physicians (Hempelmann & Krafts, 2013). These miasmatists fended off theoretical challenges by those who believed that malaria was caused by dramatic temperature changes and a failure of the human body to adapt (Oldham, 1871), or rainfall disrupting the algae in standing water (Gagliardi, 1870) – which seem remarkably similar to persistently held beliefs in Tanzania (Gessler, et al., 1995; Winch, et al., 1997; Mboera, et al., 2010) – and spent time humorously examining the ancient views that disease was spread by metaphysical or supernatural forces (Barker, 1863).

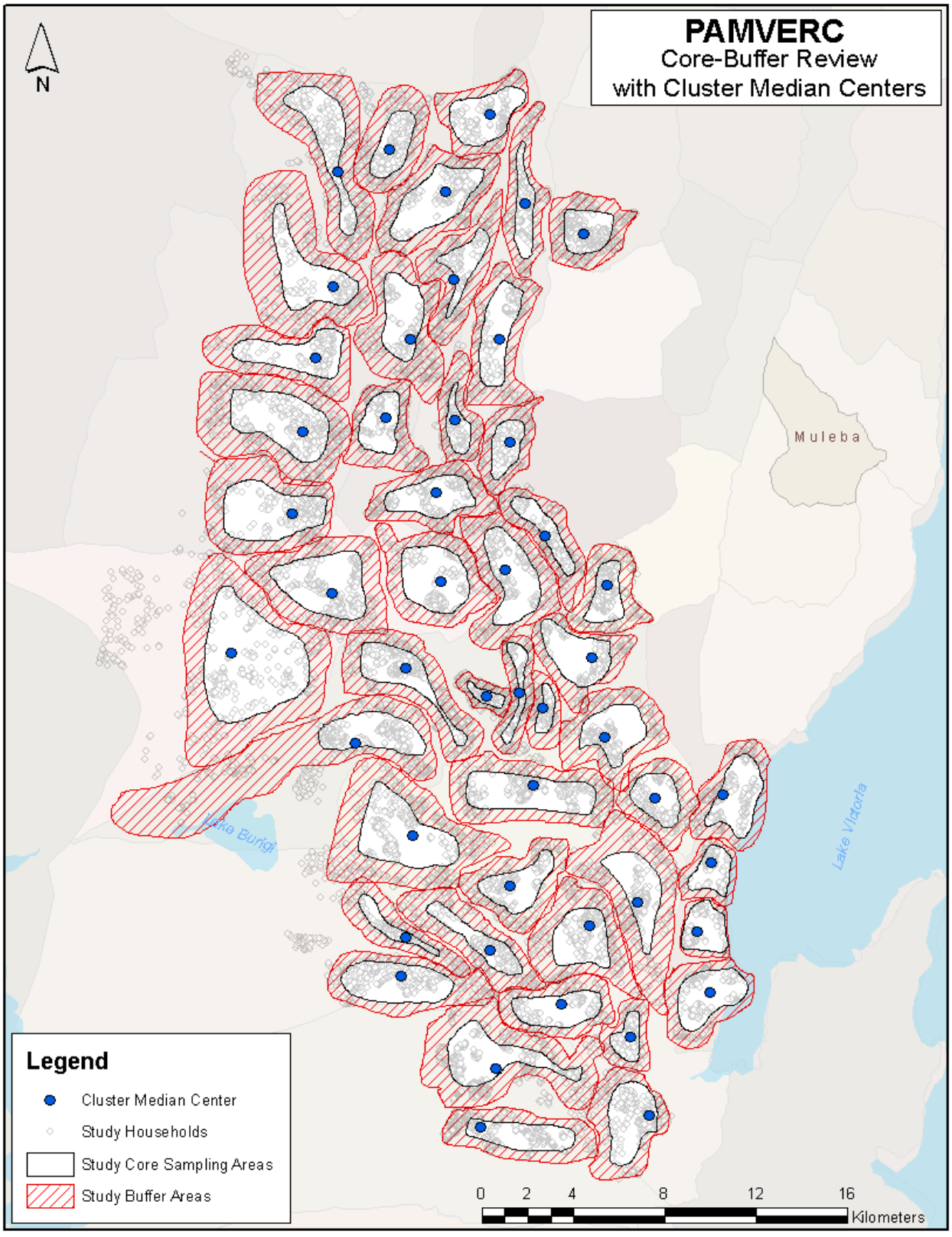
Indeed, the degree of certainty with which the proponents of these theories present their evidence should give pause to any scholar that purports to fully understand the causes and risk factors surrounding the disease. As we strive to develop a clear understanding of the

transmission of malaria through the factors associated with its causes, we must remember that these “truths” we hold as evident in modern science may be unceremoniously disrupted by the next Antonie van Leeuwenhoek or Robert Koch as were the theories posed by the miasmatists. But while these historical theories were almost unexceptionally incorrect from a modern lens, they all shared a uniform conclusion which remains at the heart of this field of study: “the place in which it is produced, or that which is nearest, ought to suffer most from its action” (MacCulloch, 1827). It is in this that spatial epidemiology places its faith: spatial associations will define the gaps in our knowledge, and their analysis will continue to push forward our understanding of the factors surrounding the propagation of one of the world’s most devastating plagues such that one day it may be brought to an end.

Table 6.1 *Post-Intervention Odds of Malaria Infection (P. Falciparum or otherwise) with Distance*

	Simple Logistic		Adjusted Model Distance	
Parameter	OR (95% CI)	p value	OR (95% CI)	p value
Age (years)	1.10 (1.08-1.11)	<0.0001	1.12 (1.11-1.14)	<0.0001
Sex (Male vs Female)	1.11 (1.01-1.23)	0.0328	1.15 (1.04-1.28)	0.0070
Elevation (per 100m)	0.66 (0.62-0.70)	<0.0001	0.66 (0.63-0.70)	<0.0001
Distance from Cluster Center (per 100m)	1.03 (1.02-1.03)	<0.0001	1.02 (0.01-1.03)	<0.0001
Household has Open Eaves	1.66 (1.50-1.85)	<0.0001	1.47 (1.31-1.65)	<0.0001
Roofing Material				
Grass/Leaves & Partial Metal	1.62 (1.43-1.84)	<0.0001	1.13 (0.98-1.31)	0.0999
Metal Sheets	REF		REF	
Head of Household Schooling				
No Education	2.81 (2.07-3.80)		1.71 (1.23-2.36)	
Primary Education	2.15 (1.60-2.88)	<0.0001	1.54 (1.13-2.10)	0.0043
Secondary/Tech or Higher Education	REF		REF	
Wealth Quintile				
Poorest	1.98 (1.69-2.31)		1.47 (1.22-1.76)	
Poorer	1.56 (1.34-1.83)		1.29 (1.09-1.53)	
Middle	1.38 (1.19-1.60)	<0.0001	1.21 (1.03-1.42)	0.0012
Less Poor	1.24 (1.06-1.44)		1.15 (0.98-1.36)	
Least Poor	REF		REF	

Figure 6.1 *Review of Core and Buffer Areas for Post-Intervention Sampling Compared with Median Cluster Centers*



7. References

Alaii, J. A., Hawley, W. A., Kolczak, M. S., Ter Kuile, F. O., Gimnig, J. E., Vulule, J. M., . . . Phillips-Howard, P. A. (2003). Factors affeting use of permethrin-treated bed nets during a randomized controlled trial in western Kenya. *Journal of Tropical Medicine & Hygiene, 68*(Suppl 4), 137-141.

Alonso, P. L., Lindsay, S. W., Armstrong, J. R., Conteh, M., Hill, A. G., David, P. H., . . . Greenwood, B. M. (1991, June 22). The effect of insecticide-treated bed nets on mortality of Gambian children. *The Lancet, 337*, 1499-1502.

Alves, A. T., Nobre, F. F., & Waller, L. A. (2016). Exploring spatial patterns in the associations between local AIDS incidence and socioeconomic and demographic variables in the state of Rio de Janeiro, Brazil. *Spatial and Spatio-temporal Epidemiology, 17*, 85-93.

APHA. (2015). *Control of Communicable Diseases Manual* (20th ed.). (D. L. Heymann, Ed.) Washington: Alpha Press.

ArcGIS. (2015). ArcGIS 10.4.1. Redlands, CA, USA: Envicornmental Systems Research Institute (ESRI) Inc.

Asare, E. O., Tompkins, A. M., Amekudzi, L. K., Ermert, V., & Redl, R. (2016). Mosquito breeding site water temperature observations and simulations towards improved vector-borne models for Africa. *Geospatial Health, 11*(s1).

Astuti, H. N., Saputro, D. R., & Susanti, Y. (2017). Mixed geographically weighted regression (MGWR) model with weighted adaptive bi-square for case of dengue hemorrhagic fever (DHF) in Surakarta. *Journal of Physics: Conference Series, 885*.

Atieli, H. E., Zhou, G., Afrane, Y., Lee, M., Mwanzo, I., Githeko, A. K., & Yan, G. (2011). Insecticide-treated net (ITN) ownership, usage, and malaria transmission in the highlands of western Kenya. *Parasites & Vectors, 4*(113), 1-10.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical Modelling and Analysis for Spatial Data.* New York: Chapman & Hall.

Bannister-Tyrrell, M., Xa, N. X., Kattenberg, J. H., Van, N. V., Dung, V. K., Hieu, T. M., . . . Erhart, A. (2018). Micro-epidemiology of malaria in an elimination setting in Central Vietnam. *Malaria Journal, 17*(119), 1-15.

Barker, T. H. (1863). *Malaria and Miasmata and their Influence in the Production of Typhus and Thyphoid Fevers, Cholera, and the Exanthemata.* London: John W. Davies.

Bayoh, M. N., & Lindsay, S. W. (2003). Effect of temperature on the development of the aquatic stages of Anopheles gambiae sensu stricto (Diptera: Culicidae). *Bulletin of Entomological Research, 93*, 375-381.

Bayoh, M. N., & Lindsay, S. W. (2004). Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito Anopheles gambiae in the laboratory. *Medical and Veterinary Entomology, 18*, 174-179.

Binka, F. N., & Adongo, P. (1997). Acceptability and use of insecticide impregnated bednets in northern Ghana. *Tropical Medicine and International Health, 2*(5), 499-507.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G. H., . . . Smith, P. G. (1996). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health, 1*(2), 147-154.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G. H., . . . Smith, P. G. (1998). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health, 1*(2), 147-154.

Blanford, J. I., Blanford, S., Crane, R. G., Mann, M. E., Paaijmans, K. P., Schreiber, K. V., & Thomas, M. B. (2013). Implications of temperature variation for malaria parasite development across Africa. *Scientific Reports, 3*(1300).

Bødker, R., Akida, J., Shayo, D., Kisinza, W., Msangeni, H. A., Pedersen, E. M., & Lindsay, S. W. (2003). Relationship Between Altitude and Intensity of Malaria Transmission in the Usambara Mountains, Tanzania. *Journal of Medical Entomology, 40*(5), 706-717.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8*, 3-62.

Bousema, T., Stresman, G., Baidjoe, A. Y., Bradley, J., Knight, P., Stone, W., . . . Cox, J. (2016). The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLOS Medicine, 13*(4).

Bradley, J., Rehman, A. M., Schwabe, C., Vargas, D., Monti, F., Ela, C., . . . Kleinschmidt, I. (2013). Reduced Prevalence of Malaria Infection in Children Living in Houses with Window Screening or Closed Eaves on Bioko Island, Equatorial Guinea. *PLoS ONE, 8*(11), 1-7.

Brom, A., & Natonik, A. (2017). Estimation of geotechnical parameters on the basis of geophysical methods and geostatistics. *Computational Trends in Geoscience, 6*(2), 70-79.

Brooker, S., Clarke, S., Njagi, J. K., Polack, S., Mugo, B., Estamblale, B., . . . Cox, J. (2004). Spatial clustering of malaria and associated risk factors during an epidemic in a highland area of western Kenya. *Tropical Medicine & International Health, 9*(7), 757-766.

Brunsdon, C., Aitkin, M., Fotheringham, S., & Charlton, M. (1999). A comparison of random coefficient modelling and geographically weighted regression for spatially non-stationary regression problems. *Geographical & Environmental Modelling, 3*(1), 47-62.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance test for geographically weighted regression. *Journal of Regional Science, 39*(3), 497-524.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis, 28*(4).

Bull, P. C., Lowe, B. S., Kortok, M., Molyneux, C. S., Newbold, C. I., & Marsh, K. (1998). Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nature Medicine, 4*(3), 358-360.

Burgert, C. R., Bradley, S. E., Arnold, F., & Ekert, E. (2014). Improving estimates of insecticide-treated mosquito net coverage from household surveys: Using geographic coordinates to account for endemicity. *Malaria Journal, 13*(254), 1-13.

Burt, J. E., & Barber, G. M. (1996). *Elementary Statistics for Geographers* (2nd ed.). New York: Guilford Press.

CDC. (2016, 04 13). *Malaria*. Retrieved from Centers for Disease Control and Prevention: https://www.cdc.gov/malaria/

Chilès, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty* (Second ed.). John Wiley & Sons Inc.

Christiansen-Jucht, C., Parham, P. E., Saddler, A., Koella, J. C., & Basáñez, M. (2014). Temperature during larval development and adult maintenance influences the survival of Anopheles gambiae s.s. *Parasites & Vectors, 7*(489).

Cressie, N. (1960). *Statistics for Spaital Data.* New York: Wiley.

Cressie, N. (1993). *Statistics for Spatial Data.* New York: John Wiley & Sons, Inc.

Cressie, N., & Hawkins, D. M. (1980). Robust Estimation of the Variogram: I. *Mathematical Geography, 12*(2), 115-125.

Curtis, C. F., Maxwell, C. A., Magesa, S. M., Rwegoshora, R. T., & Wilkes, T. J. (2006). Insecticide-treated bed-nets for malaria mosquito control. *Journal of the American Mosquito Control Association*, 501-506.

D'Alessandro, U., Olaleye, B., McGuire, W., Langerock, P., Aikins, M., & al, e. (1995). Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. *Lancet, 345*(8948), 479-83.

Daygena, T. Y., Massebo, F., & Lindtjørn, B. (2017). Variation in species composition and infection rates of Anopheles mosquitoes at different altitudinal transects, and the risk of malaria in the highland of Dirashe Woreda, south Ethiopia. *Parasites & Vectors, 10*(343), 1-13.

Dear, N. F., Kadangwe, C., Mzilahowa, T., Bauleni, A., Mathanga, D. P., Duster, C., . . . Wilson, M. L. (2018). Household-level and surrounding peri-domestic environmental characteristics associated with malaria vectors Anopheles arabiensis and Anopheles

funestus along an urban–rural continuum in Blantyre, Malawi. *Malaria Journal, 17*(229), 1-14.

Delrieu, I., Leboulleux, D., Ivinson, K., Gessner, B. D., Chandramohan, D., Churcher, T., . . . Smith, T. (2015). Design of a Phase III cluster randomized trial to assess the efficacy and safety of a malaria transmission blocking vaccine. *Vaccine, 33*, 1518-1526.

Deressa, W., Yihego, Y. Y., Kebede, Z., Batisso, E., Tekalegne, A., & Dagne, G. A. (2014). Effect of combining mosquito repellent and insecticide treated net on malaria prevalence in Southern Ethiopia: a cluster-randomised trial. *Paraties & Vectors, 7*(132).

Desowitz, R. S. (1991). *The Malaria Capers: More tales of parasites and people, research and reality.* New York: W. W. Norton & Company, Inc.

Devkota, M. L., Hatfield, G., & Chintala, R. (2014). Effect of Sample Size on the Performance of Ordinary Least Squares and Geographically Weighted Regression. *British Journal of Mathematics & Computer Science, 4*(1), 1-21.

ESRI. (2018, 09 06). *Median Center*. Retrieved from ArcGIS Desktop: http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/median-center.htm

Ezanno, P., Aubry-Kientz, M., Cailly, P., L'Ambert, G., Toty, C., Balenghein, T., & Tran, A. (2015). A generic weather-driven model to predict mosquito populaton dynamics applied to species of Anopheles, Culex and Ades genera of southern France. *Preventive Veterinary Medicine, 120*, 39-50.

Froeschl, G., Saathoff, E., Kroidl, I., Berens-Risha, N., Clowes, P., Maboko, L., . . . Hoelscher, M. (2018). Reduction of malaria prevalence after introduction of artemisinin-combination-therapy in Mbeya Region, Tanzania: results from a cohort study with 6773 participants. *Malaria Journal, 17*(245), 1-11.

Gagliardi, J. (1870). A Probable Cause of Malaria. *Nature, 1*, 481.

Gessler, M. C., Msuya, D. E., Nkunya, M. H., Schär, A., Heinrich, M., & Tanner, M. (1995). Traditional healers in Tanzania: the perception of malaria and its causes. *Journal of Ethnopharmacology, 48*, 119-130.

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R., Johnston, G. L., . . . Hay, S. I. (2011). A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria Journal, 10*(378), 1-16.

Ghebreyesus, T. A., Haile, M., Witten, K. H., Getachew, A., Yohannes, M., Lindsay, S. W., & Byass, P. (2000). Household risk factors for malaria among children in the Ethiopian highlands. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 94*(1), 17-21.

Govella, N. J., Okumu, F. O., & Killeen, G. F. (2010). Short Report: Insecticide-Treated Nets Can Reduce Malaria Transmission by Mosquitoes Which Feed Outdoors. *American Journal of Tropical Medicine and Hygiene, 82*(3), 415-419.

Grassi, B., Bignami, A., & Bastianelli, G. (1899). Medical zoology: Further researches upon the cycle of human malaria in the body of the mosquito. *The Indian Medical Gazette*, 104-107.

GWR4. (2015, 03). GWR4 v4.0.9. Maynooth , Ireland: National Center for Geocomputation, National University of Ireland Maynooth and Department of Geography, Ritsumeikan University, Japan.

Haque, U., Scott, L. M., Hashizume, M., FIsher, E., Haque, R., Yamamoto, T., & Glass, G. E. (2012). Modelling malaria treatment practices in Bangladesh using spatial statistics. *Malaria Journal, 11*(63), 1-11.

Hasyim, H., Nursafingi, A., Haque, U., Montag, D., Groneberg, D. A., Dhimal, M., . . . Müller, R. (2018). Spatial modelling of malaria cases associated with environmental factors in South Sumatra, Indonesia. *Malaria Journal, 17*(87), 1-15.

Hawley, W. A., Phillips-Howard, P. A., Ter Kule, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., . . . Hightower, A. W. (2003). Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. *American Journal of Tropical Medicine and Hygiene, 68*(Suppl 4), 121-127.

Heggenhougen, H. K., Hackethal, V., & Vivek, P. (2003). *The behavioural and social aspects of malaria and its control: An introduction and annotated bibliography.* UNDP/World Bank/WHO, Special Programme for Research and Training in Tropical Diseases. Geneva: World Health Organization.

Hempelmann, E., & Krafts, K. (2013). Bad air, amulets and mosquitoes: 2,000 years of changing perspectives on malaria. *Malaria Journal, 12*(232), 1-13.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis.* New York: Wiley.

Homan, T., Marie, N., Hiscox, A., Di Pasquale, A., Kiche, I., Onoka, K., . . . Takken, W. (2016). Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: an explorative study. *Malaria Journal, 15*.

Homenauth, E. (2016). *Evaluating multi-level risk factors for malaria and arboviral infections in regions of Tanzania.* Thesis, University of Ottawa, School of Epidemiology, Public Health and Preventative Medicine, Ottawa.

Homenauth, E., Kajeguka, D., & Kulkarni, M. A. (2017). Principal component analysis of socioeconomic factors and their association with malaria and arbovirus risk in Tanzania: a sensitivity analysis. *Socioeconomic Factors and Health, 71*, 1046-1051.

Houweling, T. A., Kunst, A. E., & Mackenbach, J. P. (2003). Measuring health inequality among children in developing countries: does the choice of the indicator of economic status matter? *International Journal for Equity in Health, 2*(8), 1-12.

Howard, S. C., Omumbo, J., Nevill, C., Some, E. S., Donnelly, C. A., & Snow, R. W. (2000). Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of The Royal Society of Tropical Medicine and Hygiene, 94*(4), 357-360.

Howe, L. D., Hargreaves, J. R., & Huttly, S. R. (2008). Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerging Themes in Epidemiology, 5*(3), 1-14.

Huque, H., Bondell, H. D., Carroll, R. J., & Ryan, L. M. (2016). Spatial regression with covariate measurement error: A semiparametric approach. *Biometrics, 72*, 678-686.

IHME. (2016). *GBD Compare Data Visualization*. (University of Washington) Retrieved 01 24, 2017, from Institute for Health Metrics and Evaluation: http://vizhub.healthdata.org/gbd-compare

IVCC. (2018, 09 09). *Why do we need 3 new Insecticides?* Retrieved from Innovative Vector Control Consortium: http://www.ivcc.com/creating-solutions/why-do-we-need-3-new-insecticides

Jaffer, A., Protopopoff, N., Mosha, F. W., Malone, D., Rowland, M. W., & Oxborough, R. M. (2015). Evaluating the sterilizing effect of pyriproxyfen treated mosquito nets against Anopheles gambiae at different blood-feeding intervals. *Acta Tropica, 150*, 131-135.

Jarvis, C., Di Tanna, G., Lewis, D., Alexander, N., & Edmunds, W. (2017). Spatial analysis of cluster randomised trials: a systematic review of analysis methods. *Emerging Themes in Epidemiology, 14*(12).

Jian, X., Olea, R. A., & Yu, Y. (1996). Semivariogram modeling by weighted least squares. *Computers & Geosciences, 22*, 387-397.

Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics.* New York: Academic Press.

Jung, I., Kulldorff, M., & Klassen, A. (2007). A spatial scan statistic for ordinal data. *Statistics in Medicine, 26*, 1594-1607.

Kajla, M., Bhattacharya, K., Gupta, K., Bannerjee, U., Kakani, P., Gupta, L., & Kumar, S. (2016). Identification of the Temperature Induced Larvicidal Efficacy of Agave angustifolia against Aedes, Culex, and Anopheles Larvae. *Frontiers in Public Health, 3*.

Kamol-Ratanakul, P., & Prasittisuk, C. (1992). The effectiveness of permethrin-impregnated bed nets against malaria for migrant workers in eastern Thailand. *American Journal of Tropical Medicine and Hygiene, 47*(3), 305-9.

Kawada, H., Ohashi, K., Dida, G. O., Sonye, G., Njenga, S. M., Mwandawiro, C., & Minakawa, N. (2014). Insecticidal and repellent activities of pyrethroids to the three major pyrethroid-resistant malaria vectors in western Kenya. *Parasites & Vectors, 7*(208).

Kelly-Hope, L. A., Hemingway, J., & McKenzie, F. E. (2009). Environmental factors associated with the malaria vectors Anopheles gambiae and Anopheles funestus in Kenya. *Malaria Journal, 8*(268), 1-8.

Kiefer, J. (1953). Sequential minmax search for a maximum. *Proceedings of the American Mathematical Society, 4*, 502-506.

Killeen, G. F., Smith, T. A., Ferguson, H. M., Mshinda, H., Abdulla, S., Lengeler, C., & Kachur, S. P. (2007). Preventing Childhood Malaria in Africa by Protecting Adults from Mosquitoes with Insecticide-Treated Nets. *PLoS Medicine, 4*(7), 1246-1258.

Kleinschmidt, I., Sharp, B. L., Clarke, G. P., Curtis, B., & Fraser, C. (2001). Use of Generalized Linear Mixed Models in the Spatial Analysis of Small-Area Malaria Incidence Rates in KwaZulu Natal, South Africa. *American Journal of Epidemiology, 153*(12), 1213-1221.

Kristan, M., Abeku, T. A., & Lines, J. (2018). Effect of environmental variables and kdr resistance genotype on survival probability and infection rates in Anopheles gambiae (s.s.). *Parasites & Vectors, 11*(560), 1-10.

Kroeger, A. (1995). Insecticideimpregnated bed nets for malaria control: varying experiences from Ecuador, Colombia, and Peru concerning acceptability and effectiveness. *American Journal of Tropical Medicine and Hygiene, 53*(4), 313-23.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods, 26*, 1481-1496.

Kulldorff, M., Huang, L., & Konty, K. (2009). A spatial scan statistic for continuous data based on the normal probability model. *International Journal of Health Geographics, 8*(58), 1-9.

Lamorey, G., & Jacobson, E. (1995). Estimation of Semivariogram Parameters and Evaluation of the Effects of Data Sparsity. *Mathematical Geology, 27*(3), 327-358.

Lengeler, C. (2009). Insecticide-treated bed nets and curtains for perventing malaria. (C. I. Group, Ed.) *Cochrane Database of Systematic Reviews*(2), pp. 1-55.

Lenhart, A., Orleus, N., Maskill, R., Alexander, N., Streit, T., & McCall, P. J. (2008). Insecticide-treated bednets to control dengue vectors: preliminary evidence from a controlled trial in Haiti. *Tropical Medicine and International Health, 13*(1), 56-67.

Lindsay, S., Snow, R., Broomfield, G., Janneh, M., Wirtz, R., & Greenwood, B. (1989). Impact of permethrin-treated bednets on malaria transmission by the Anopheles gambiae complex in the Gambia. *Medical and Veterinary Entomology, 3*(4), 263-71.

Luxemburger, C., Perea, W., Delmas, G., Pruja, C., Pecoul, B., & Moren, A. (1994). Permethrin-impregnated bed nets for the prevention of malaria in schoolchildren on the ThaiBurmese border. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 88*(2), 155-9.

Lymio, E. O., Takken, W., & Koella, J. C. (1992). Effect of rearing temperature and larval density on larval survival, age at pupation and adult size of Anopheles gambiae. *Entomologia Experimentalis et Applicata, 63*, 265-271.

Lyons, C. L., Coetzee, M., & Chown, S. L. (2013). Stable and fluctuating temperature effects on the development rate and survival of two malaria vectors, Anopheles arabiensis and Anopheles funestus. *Parasites & Vectors, 6*(104), 1-9.

MacCulloch, J. (1827). *Malaria: An Essay on the Production and Propagation of this Poison, and on the Nature and Localities of the Places by which it is Produced: with an Enumeration of the Diseases Caused by it, and of the Means of Preventing or Diminishing Them.* London: Longman, Rees, Orme, Brown, and Green.

Manson, P. (1900, September 29). Experimental proof of the mosquito-malaria theory. *The Lancet*, 923-925.

Manyangadze, T., Chimbari, M. J., Macherera, M., & Mukaratirwa, S. (2017). Micro-spatial distribution of malaria cases and control strategies at ward level in Gwanda district, Matabeleland South, Zimbabwe. *Malaria Journal, 16*(476), 1-11.

Marsh, K. (1992). Malaria - a neglected disease? *Parasitology, 104*, S53-S69.

Mayfield, H. J., Lowry, J. H., Watson, C. H., Kama, M., Nilles, E. J., & Lau, C. L. (2018). Use of geographically weighted logistic regression to quantify spatial variation in the environmental and sociodemographic drivers of leptospirosis in Fiji: a modelling study. *Lancet Planetary Health, 2*, e223-32.

Maziarz, M., Nabalende, H., Otim, I., Legason, I. D., Kinyera, T., Ogwang, M. D., . . . Mbulaiteye, S. M. (2018). A cross-sectional study of asymptomatic Plasmodium falciparum infection burden and risk factors in general population children in 12 villages in northern Uganda. *Malaria Journal, 17*(240), 1-15.

Mboera, L. E., Shayo, E. H., Senkoro, K. P., Rumisha, S. F., Mlozi, M. R., & Mayala, B. K. (2010). Knowledge, perceptions and practices of farming communities on linkages between malaria and agriculture in Mvomero District, Tanzania. *Acta Tropica, 113*, 139-144.

Mei, C., He, S., & Fang, K. (2004). A note on the mixed geographically weighted regression model. *Journal of Regional Science, 44*(1), 143-157.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics, 21*(6), 1087-1092.

Midekisa, A., Beyene, B., Mihretie, A., Bayabil, E., & Wimberly, M. C. (2015). Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. *Parasites & Vectors, 8*(339), 1-11.

MoHCDGEC, Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS), & 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.

Mordecai, E. A., Paaijmans, K. P., Johnson, L. R., Balzer, C., Ben-Horin, T., Moor, E., . . . Lafferty, K. D. (2013). Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecology Letters, 16*, 22-30.

Mtove, G., Mugasa, J. P., Messenger, L. A., Mangesho, P., Plucinski, M., Hashimu, R., . . . Kisinza, W. N. (2016). The effectiveness of non-pyrethroid insecticide-treated durable wall lining to control malaria in rural Tanzania: study protocol for a two-armed cluster randomized trial. *BMC Public Health, 16*(633).

Mwesigwa, J., Achan, J., Di Tanna, G. L., Affara, M., Jawara, M., Worwui, A., . . . D'Alessandro, U. (2017). Residual malaria transmission dynamics varies across The Gambia despite high coverage of control interventions. *PloS One, 12*(11), 1-24.

Nakaya, T. (2016). *GWR4.09 User Manual.* GWR 4 Development Team.

Nakaya, T., Fotheringham, A. S., Charlton, M., & Brundson, C. (2009). Semiparametric geographically weighted generalised linear modelling in GWR 4.0.

National Bureau of Statistics. (2016). *Basic Demographic and Socio-Economic Profile Kagera Region.* Ministry of Finance, National Bureau of Statistics. Dar es Salaam: The United Republic of Tanzania.

Ndiath, M. M., Cisse, B., Ndiaye, J. L., Gomis, J. F., Bathiery, O., Dia, A. T., . . . Faye, B. (2015). Application of geographically-weighted regression analysis to assess risk factors for malaria hotspots in Keur Soce health and demographic surveillance site. *Malaria Journal, 14*(463).

Nevil, C. G., Some, E. S., Mung'ala, V. O., Mutemi, W., New, L., Marsh, K., . . . Snow, R. (1996). Insecticide-treated bednets reduce child mortality and severe morbidity from malaria among children on the Kenyan coast. *Tropical Medicine and International Health, 1*(2), 139-146.

Nicholson, M. C., & Mather, T. N. (1996). Methods for Evaluating Lyme Disease Risks Using Geographic Information Systems and Geospatial Analysis. *Journal of Medical Entomology, 33*(5), 711-720.

NMCP. (2014). *National Malaria Strategic Plan 2014-2020.* Ministry of Health and Social Welfare. Tanzania.

NPIC. (2017). *Piperonyl Butoxide General Fact Sheet.* Oregon State University, National Pesticide Information Center. Retrieved from npic.orst.edu/factsheets/pbogen.html

Nusinovici, S., Frössling, J., Widgren, S., Beaudeau, F., & Lindberg, A. (2015). Q fever infection in dairy cattle herds: increased risk with high wind speed and low precipitation. *Epidemiological Infection*, 1-11.

Octavianty, Jaya, I. G., & Toharudin, T. (2017). Geographically weighted poisson regression semiparametric on modeling of the number of tuberculosis cases (Case study: Bandung city). *AIP Conference Proceedings, 1827*, 1-7.

Oldham, C. F. (1871). *What is Malaria? And why is it most intense in hot climates?* London: H. K. Lewis.

Olea, R. A. (1974). Optimal contour mapping using Kriging. *Journal of Geophysical Research, 79*(5), 695-702.

Olea, R. A. (2006). A six-step practical approach to semivariogram modeling. *Stochastic Environmental Research and Risk Assessment, 20*(5), 307-318.

Oliver, M. A., & Webster, R. (1986). Semi-variograms for modelling the spatial pattern of landform and soil properties. *Earth Surface Processes and Landforms, 11*, 491-504.

Oritz, B. V., Sullivan, D. G., Perry, C., & Vellidis, G. (2007). Spatial Variability of Root Knot Nematodes in Relation to Within Field Variability of Soil Properties. *ASABE Annual International Meeting.* American Society of Agricultural and Biological Engineers.

Ouedraogo, B., Inoue, Y., Kambiré, A., Sallah, K., Dieng, S., Tine, R., . . . Gaudart, J. (2018). Spatio-temporal dynamic of malaria in Ouagadougou, Burkina Faso, 2011–2015. *Malaria Journal, 17*(138), 1-12.

Paaijmans, K. P., Blanford, S., Chen, B. H., & Thomas, M. B. (2012). Warmer temperatures reduce the vectoral capacity of malaria mosquitoes. *Biology Letters, 8*, 465-468.

Pakdad, K., Hanafi-Bojd, A. A., Vatandoost, H., Sedaghat, M. M., Raeisi, A., Moghaddam, A. S., & Roroushani, A. R. (2017). Predicting the potential distribution of main malaria vectors Anopheles stephensi, An. culicifacies s.l. and An. fluviatilis s.l. in Iran based on maximum entropy model. *Acta Tropica, 169*, 93-99.

PAMVERC. (2014). *Detailed Work Plan and Protocol (10 August 2014) v1.6: Evaluation of a novel long lasting insecticidal net and indoor residual spray product, separately and together, against malaria transmitted by pyrethroid resistant mosquitoes.* Pan-African Malaria Vector Research Consortium.

Pardanani, N., & Mather, T. N. (2004). Lack of Spatial Autocorrelation in Fine-Scale Distributions of Ixodes scapularis (Acari: Ixodidae). *Journal of Medical Entomology, 41*(5), 861-864.

PMI. (2016). *Malaria Operational Plan FY 2016.* President's Malaria Initiative. Tanzania.

Protopopoff, N., Bortel, W. V., Speybroeck, N., Geertruyden, J. V., Baza, D., D'Alessandro, U., & Coosemans, M. (2009). Ranking Malaria Risk Factors to Guide Malaria Control Efforts in African Highlands. *PLoS One, 4*(11).

Protopopoff, N., Mosha, J. F., Lukaloe, E., Charlwood, J. D., Wright, A., Mwalimu, C. D., . . . Rowland, M. (2018). Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two fact. *Lancet, 391*, 1577-88.

Raghavendra, K., Chourasia, M. K., Swain, D. K., Bhatt, R. M., Uragayala, S., Dutta, G. D., & Kleinschmidt, I. (2017). Monitoring of long-lasting insecticidal nets (LLINs) coverage versus utilization: A community-based survey in malaria endemic villages of Central India. *Malaria Journal, 16*(467), 1-8.

Rao, R. (1964). The Use and Interpretation of Principal Component Analysis in Applied Research. *Sankhyā: The Indian Journal of Statistics, 26*(4), 329-358.

Roberts, D., & Matthews, G. (2016). Risk factors of malaria in children under the age of five years old in Uganda. *Malaria Journal*(15).

Rusk, A., Highfield, L., Wilkerson, J. M., Harrell, M., Obala, A., & Amick, B. (2016). Geographically-weighted regression of knowledge and behaviour determinants to anti-

malarial recommending and dispensing practice among medicine retailers in western Kenya: capacitating targeted interventions. *Malaria Journal, 15*(562).

SAS. (2012). SAS v9.4. Cary, NC, USA: SAS Institute Inc.

SAS. (2014). *SAS/STAT® 13.2 User's Guide.* Cary, NC, USA: SAS Institute Inc. Retrieved from SAS/STAT(R) 13.2 User's Guide.

SaTScan. (2016, 08). SaTScan v9.4.4. Calverton, Maryland: Harvard Medical School, Boston and Information Management Services, Inc. Retrieved from http://www.satscan.org/

Sena, L., Deressa, W., & Ali, A. (2015). Correlation of climate variability and malaria: A retrospective comparative study, southwest Ethiopia. *Ethiopian Journal of Health Science, 25*(2), 129-138.

Silkey, M., Homan, T., Marie, N., Hiscox, A., Mukabana, R., Takken, W., & Smith, T. A. (2016). Design of trials for interrupting the transmission of endemic pathogens. *Trials, 17*(278).

Singer, J., & Ovaskainen, O. (2017). Detecting the influence of environmental covariates on animal movement: a semivariance approach. *Methods in Ecology and Evolution, 8*, 561-570.

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., . . . Hay, S. I. (2012). A global map of dominant malaria vectors. *Parasites & Vectors, 5*(69), 1-11.

Sluydts, V., Durnez, L., Heng, S., Gryseels, C., Canier, L., Kim, S., . . . Coosemans, M. (2016). Efficacy of topical mosquito repellent (picaridin) plus long-lasting insecticidal nets versus long-lasting insecticidal nets alone for control of malaria: a cluster randomised controlled trial. *Lancet Infectious Diseases, 16*(10).

Snow, R., Rowan, K., Lindsay, S., & Greenwood, B. (1988). A trial of bed nets (mosquito nets) as a malaria control in a rural area of The Gambia, West Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 82*(2), 212-5.

Solomon, T., Loha, E., Deressa, W., Balkew, M., Gari, T., Overgaard, H. J., & Lindtjørn, B. (2018). Bed nets used to protect against malaria do not last long in a semi-arid area of Ethiopia: a cohort study. *Malaria Journal, 17*(239), 1-14.

Thomson, M. C., Mason, S. J., Phindela, T., & Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. *Journal of Tropical Medicine & Hygiene, 73*(1), 214-221.

Tobler, W. R. (1970, June). A computer movie simulating urban growth in the detroit region. *Economic Geography, 46*(Suppl. Proceedings. International Geographical Union, Commission on Quantitative Methods), 234-240.

Tonye, S. G., Kouambeng, C., Wounang, R., & Vounatsou, P. (2018). Challenges of DHS and MIS to capture the entire pattern of malaria parasite risk and intervention efects in countries with diferent ecological zones: the case of Cameroon. *Malaria Journal, 17*(156), 1-14.

United Nations. (2000). *55/2. United Nations Millennium Declaration.* General Assembly. United Nations.

United Nations. (2015, September 25). *Sustainable Development Goals*. (UN Web Services Section, Department of Public Information) Retrieved from United Nations: http://www.un.org/sustainabledevelopment/sustainable-development-goals/

Upadhyayula, S. M., Mutheneni, S. R., Chenna, S., Parasaram, V., & Kadiri, M. R. (2015). Climate Drivers on Malaria Transmission in Arunachal Pradesh, India. *PLoS ONE, 10*(3), 1-17.

VCAG. (2017). *How to design vector control efficacy trials: Guidance on phase III vector control field trial design provided by the Vector Control Advisory Group.* Vector Control Advisory Group. Geneva: World Health Organization.

Vyas, S., & Kumaranayake, L. (2006). *How to do (or not to do) ... Constructing socio-economic status indices: how to use principal components analysis.* London School of Hygiene and Tropical Medicine, HIVTools Research Group, Health Policy Unit, Department of Public Health and Policy. London: Oxford University Press.

Walldorf, J. A., Cohee, L. M., Coalson, J. E., Bauleni, A., Nkanauena, K., Kapito-Tembo, A., . . . Laufer, M. K. (2015). School-age children are a resevoir of malaria infection in Malawi. *PLoS ONE, 10*(7), 1-13.

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Eurpoean Journal of Soil Science, 43*, 177-192.

Were, V., Buff, A. M., Desai, M., Kariuki, S., Samuels, A., ter Kuile, F. O., . . . Niessen, L. (2018). Socioeconomic health inequality in malaria indicators in rural western Kenya: evidence from a household malaria survey on burden and care-seeking behaviour. *Malaria Journal, 17*(166), 1-10.

West, P. A., Protopopoff, N., Rowland, M., Cumming, E., Rand, A., Drakeley, C., . . . Kleinschmidt, I. (2013). Malaria Risk Factors in North West Tanzania: The Effect of Spraying, Nets and Wealth. *PLOS One, 8*(6), 1-12.

West, P. A., Protopopoff, N., Wright, A., Kivaju, Z., Tigererwa, R., Mosha, F. W., . . . Kleinschmidt, I. (2014). Indoor Residual Spraying in Combination with Insecticide-Treated Nets Compared to Insecticide-Treated Nets Alone for Protection against Malaria: A Cluster Randomized Trial in Tanzania. *PLOS Medicine, 11*(4).

WHA. (1955). *Malaria Eradication.* Committee on Programme and Budget. World Health Assembly.

WHO. (2005). *Malaria control today: Current WHO recommendations.* Roll Back Malaria. Geneva: World Health Organization.

WHO. (2007). *Insecticide-treated mosquito nets: a position statement.* Global Malaria Programme. World Health Organization.

WHO. (2013). *Report of the sixteenth WHOPES working group meeting.* Geneva: World Health Organization.

WHO. (2015). *Guidelines for the Treatment of Malaria.* Global Malaria Programme. Geneva: World Health Organization.

WHO. (2016). *Implications of insecticide resistance for malaria vector control.* Global Malaria Programme. Geneva: World Health Organization.

WHO. (2016). *World Malaria Report.* Geneva: World Health Organization.

WHO. (2017). *Malaria vector control policy recommendations and their applicability to product evaluation.* Global Malaria Programme. Geneva: World Health Organization.

WHO. (2017). *World Malaria Report.* Geneva: World Health Organization.

WHO. (2018). *Proposed evidence review group on the community effect of insecticide treated nets.* World Health Organization, Malaria Policy Advisory Committee Meeting. Geneva: World Health Organization.

WHO. (2018). *World Malaria Report 2018.* Geneva: World Health Organization.

Wiebe, A., Longbottom, J., Gleave, K., Shearer, F. M., Sinka, M. E., Massey, N. C., . . . Moyes, C. L. (2017). Geographical distributions of African malaria vector sibling species and evidence for insecticide resistance. *Malaria Journal, 16*(85), 1-10.

Winch, P. J., Makemba, A. M., Makame, V. R., Mfaume, M. S., Lynch, M. C., Premji, Z., . . . Shiff, C. J. (1997). Social and cultural factors affecting rates of regular retreatment of mosquito nets with insecticide resistance in Bagamoyo District, Tanzania. *Tropical Medicine and International Health, 2*(8), 760-770.

Yu, W., Xu, W., & Zhang, G. (2012). Using Logistic Regression and GIS to Analyze the Relationship between Precipitation and Debris Flow in Sichuan, China. *Procedia Environmental Sciences, 12*, 598-603.

Zaim, M., Ghavami, M., Nazari, M., Edrissian, G., & Nateghpour, M. (1998). Cyfluthrin (EW 050)-impregnated bednets in malaria control programme in Ghassreghand (Baluchistan, Iran). *Journal of the American Mosquito Control Association, 14*(4), 421-30.

Zheng, S., Cao, C. X., Cheng, J. Q., Wu, Y. S., Xie, X., & Xu, M. (2014). Epidemiological features of hand-foot-and-mouth disease in Shenzhen, China from 2008 to 2010. *Epidemiology and Infection, 142*(8), 1751-1762.

Zhou, Y. B., Wang, Q. X., Liang, S., Gong, Y. H., Yang, M. X., Chen, Y., . . . Jian, Q. W. (2015). Geographical variations in risk factors associated with HIV infection among drug users in a prefecture in Southwest China. *Infectious Diseases of Poverty, 4*(38).

Zhou, Y. B., Wang, Q. X., Yang, M. X., Gong, Y. H., Yang, Y., Nie, S. J., . . . Jiang, Q. W. (2016). Geographical variations of risk factors associated with HCV infection in drug users in southwestern China. *Epidemiology and Infection, 144*(6), 1291-1300.

8. Appendix

Table 8.1 *Baseline Survey Household Principal Component Characteristics*

Characteristics	n	(%)	μ	(SD)
Assets				
Radio	1343	(40.84)		
Mobile Phone	1504	(66.31)		
Bicycle	1182	(52.07)		
Motorcycle	300	(13.22)		
Car/Truck	19	(0.84)		
Canoe/Boat	28	(1.23)		
Sewing Machine	71	(3.13)		
Household				
Number of Rooms			4.9	(1.4)
Plastered Walls	247	(11.01)		
Ceiling	64	(2.82)		
Electricity	53	(2.34)		
Farmland	2245	(99.03)		
Livestock	1619	(71.32)		

Table 8.2 *Baseline Survey Principal Components Analysis Comparison of Indices*

Indicators	Index 1	Index 2	Index 3
Radio	0.472548	0.362309	0.362309
Mobile Phone	0.465108	0.359280	0.359280
Bicycle	0.431918	0.328837	0.328837
Motorcycle	0.375005	0.333571	0.333571
Car/Truck	0.229055	0.189583	0.189583
Canoe/Boat	0.114347	0.066797	0.066797
Farmland	0.036623	0.017294	0.017294
Livestock	0.326588	0.240402	0.240402
Electricity	0.244228	0.248632	0.248632
Sewing Machine		0.152526	0.152526
Number of Rooms		0.365528	0.365528
Walls Plastered			0.354282
Ceiling			0.278466
Variance Explained (%)	**22.67**	**20.20**	**20.79**

Figure 8.1 *Baseline Principal Component Index 1 Histogram*

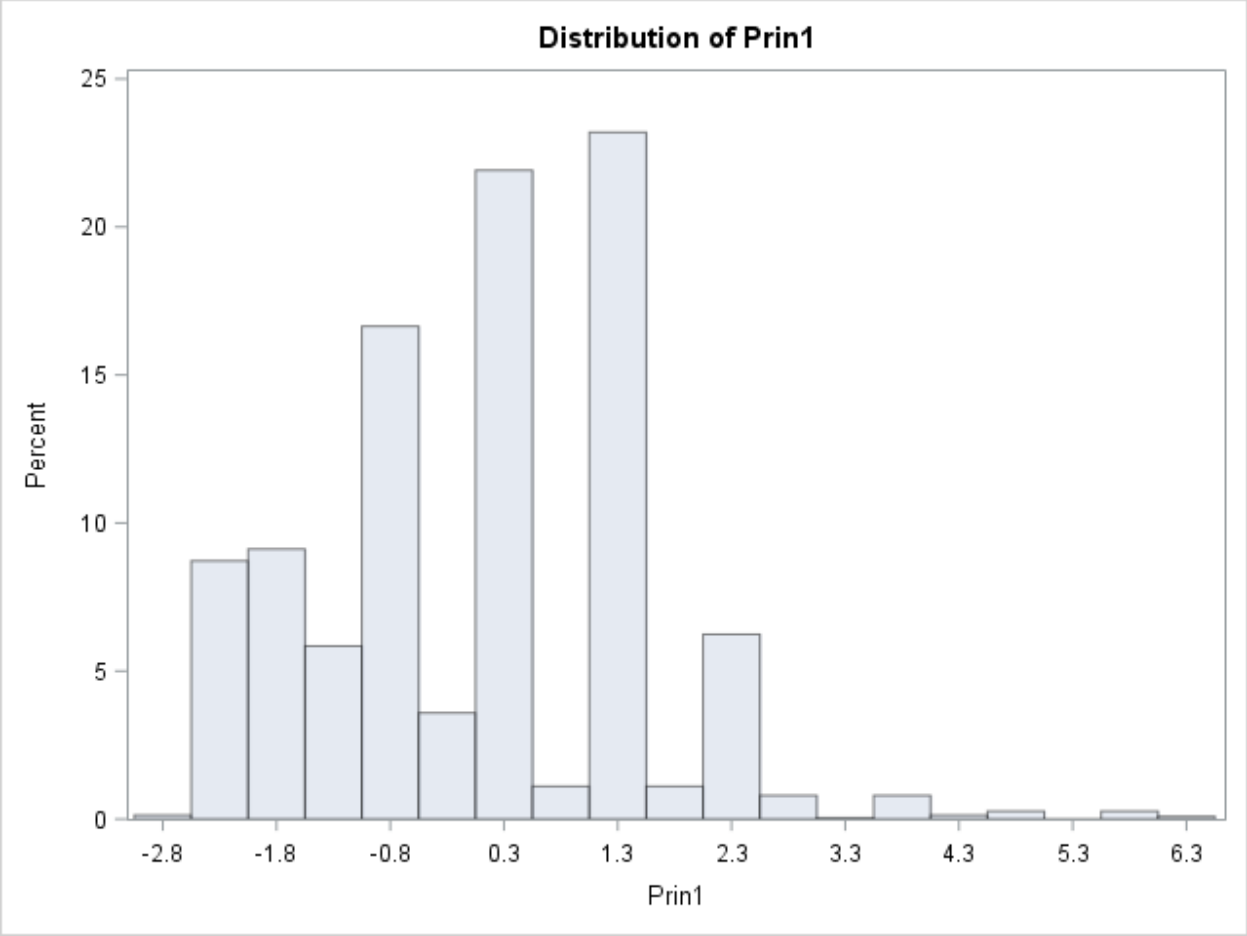


Table 8.3 *Baseline Principal Component Index 1 Defined Wealth Quintiles*

Index 1				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	452	20.01	452	20.01
1	462	20.45	914	40.46
2	472	20.89	1386	61.35
3	588	26.03	1974	87.38
4	285	12.62	2259	100.00
Frequency Missing = 621				

Figure 8.2 *Baseline Principal Component Index 2 Histogram*

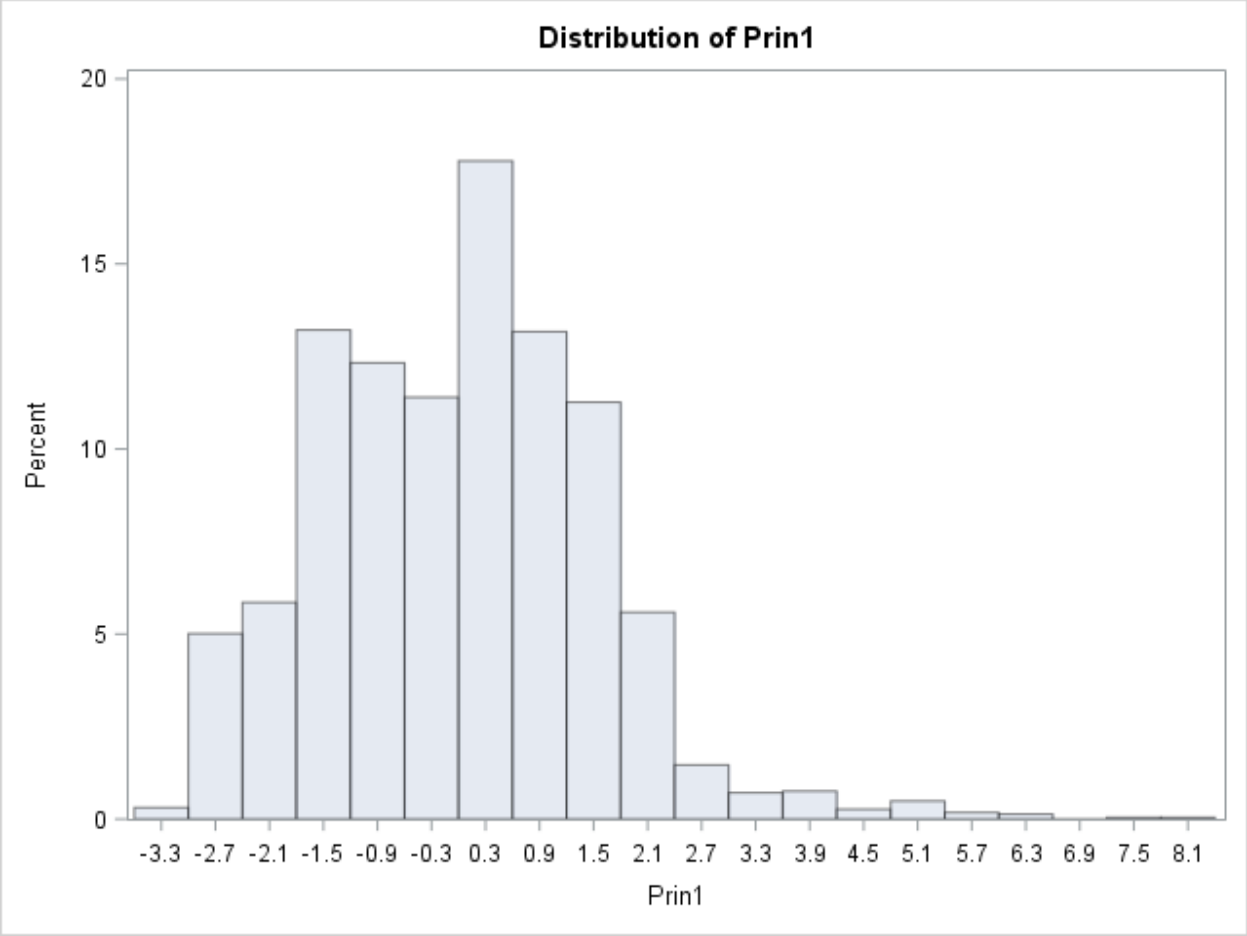


Table 8.4 *Baseline Principal Component Index 2 Defined Wealth Quintiles*

Index 2				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	451	20.00	451	20.00
1	443	19.65	894	39.65
2	473	20.98	1367	60.62
3	418	18.54	1785	79.16
4	470	20.84	2255	100.00
Frequency Missing = 625				

Figure 8.3 *Baseline Principal Component Index 3 Histogram*

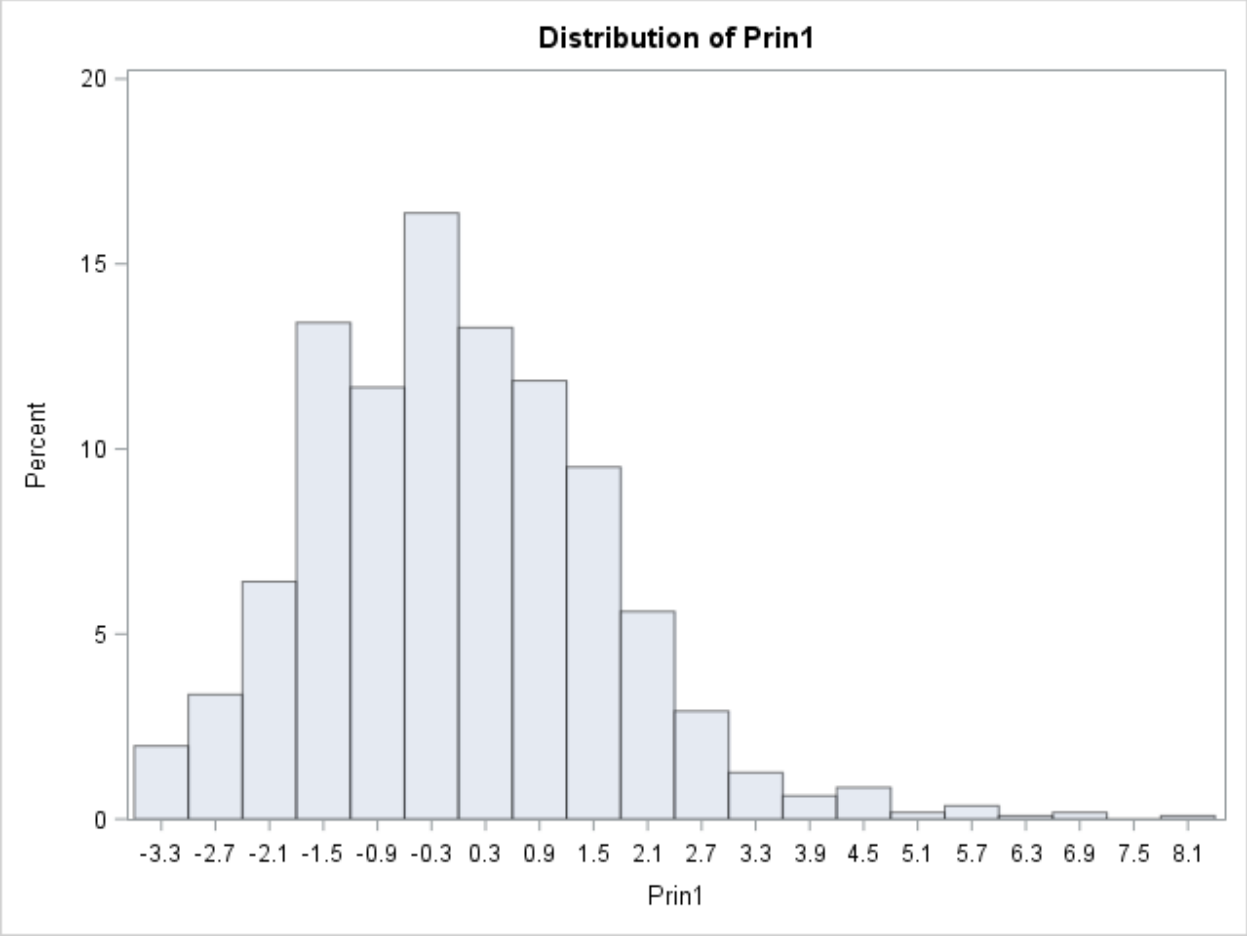


Table 8.5 *Baseline Principal Component Index 3 Defined Wealth Quintiles*

Index 3				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	441	19.78	441	19.78
1	449	20.14	890	39.93
2	446	20.01	1336	59.94
3	433	19.43	1769	79.36
4	460	20.64	2229	100.00
Frequency Missing = 651				

Table 8.6 *Post-Intervention Survey Household Principal Component Characteristics*

Characteristics	n	(%)	μ	(SD)
Assets				
Radio	1906	(57.62)		
Mobile Phone	2193	(66.31)		
Bicycle	1648	(49.85)		
Motorcycle	473	(14.31)		
Car/Truck	26	(0.79)		
Canoe/Boat	34	(1.03)		
Sewing Machine	86	(2.60)		
Household				
Number of Rooms			5.0	(1.5)
Plastered Walls	391	(11.91)		
Ceiling	49	(1.48)		
Electricity	83	(2.51)		
Farmland	3238	(98.06)		
Livestock	2284	(69.09)		

Table 8.7 *Post-Intervention Principal Components Analysis Comparison of Indices*

Indicators	Index 1	Index 2	Index 3
Radio	0.480414	0.41417	0.364027
Mobile Phone	0.473348	0.414659	0.375706
Bicycle	0.432877	0.372048	0.316559
Motorcycle	0.378695	0.360296	0.34035
Car/Truck	0.197692	0.19331	0.184136
Canoe/Boat	0.085265	0.071711	0.068813
Farmland	0.090931	0.097216	0.083128
Livestock	0.319593	0.28239	0.235355
Electricity	0.239982	0.255036	0.250637
Sewing Machine		0.190913	0.168438
Number of Rooms		0.393816	0.388485
Walls Plastered			0.340463
Ceiling			-0.226523
Variance Explained (%)	**21.69**	**20.21**	**19.11**

Figure 8.4 *Post-Intervention Principal Component Index 1 Histogram*

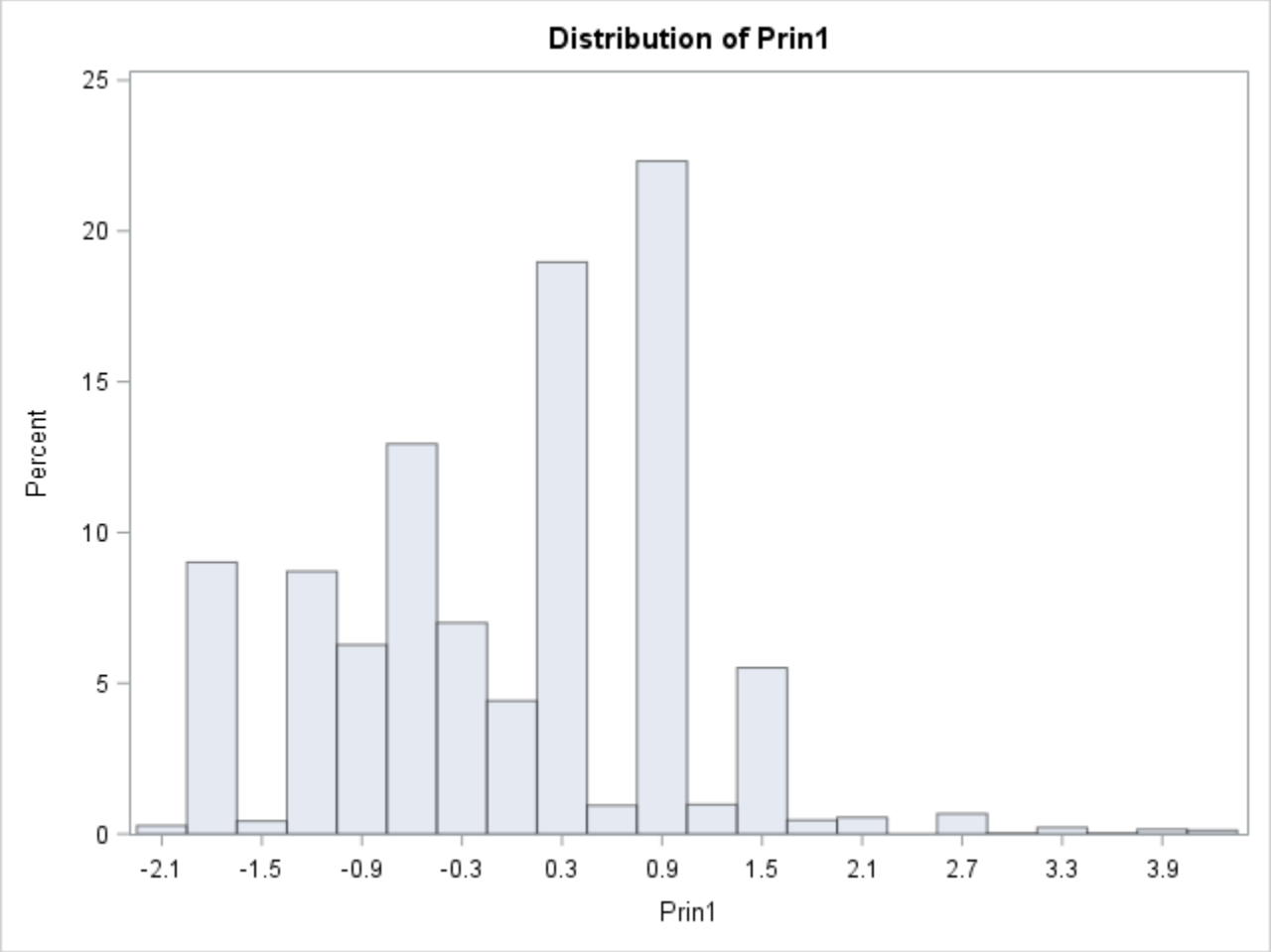


Table 8.8 *Post-Intervention Survey Principal Component Index 1 Defined Wealth Quintiles*

Index 1				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	674	20.52	674	20.52
1	614	18.7	1288	39.22
2	748	22.78	2036	62
3	853	25.97	2889	87.97
4	395	12.03	3284	100
Frequency Missing = 24				

Figure 8.5 *Post-Intervention Principal Component Index 2 Histogram*

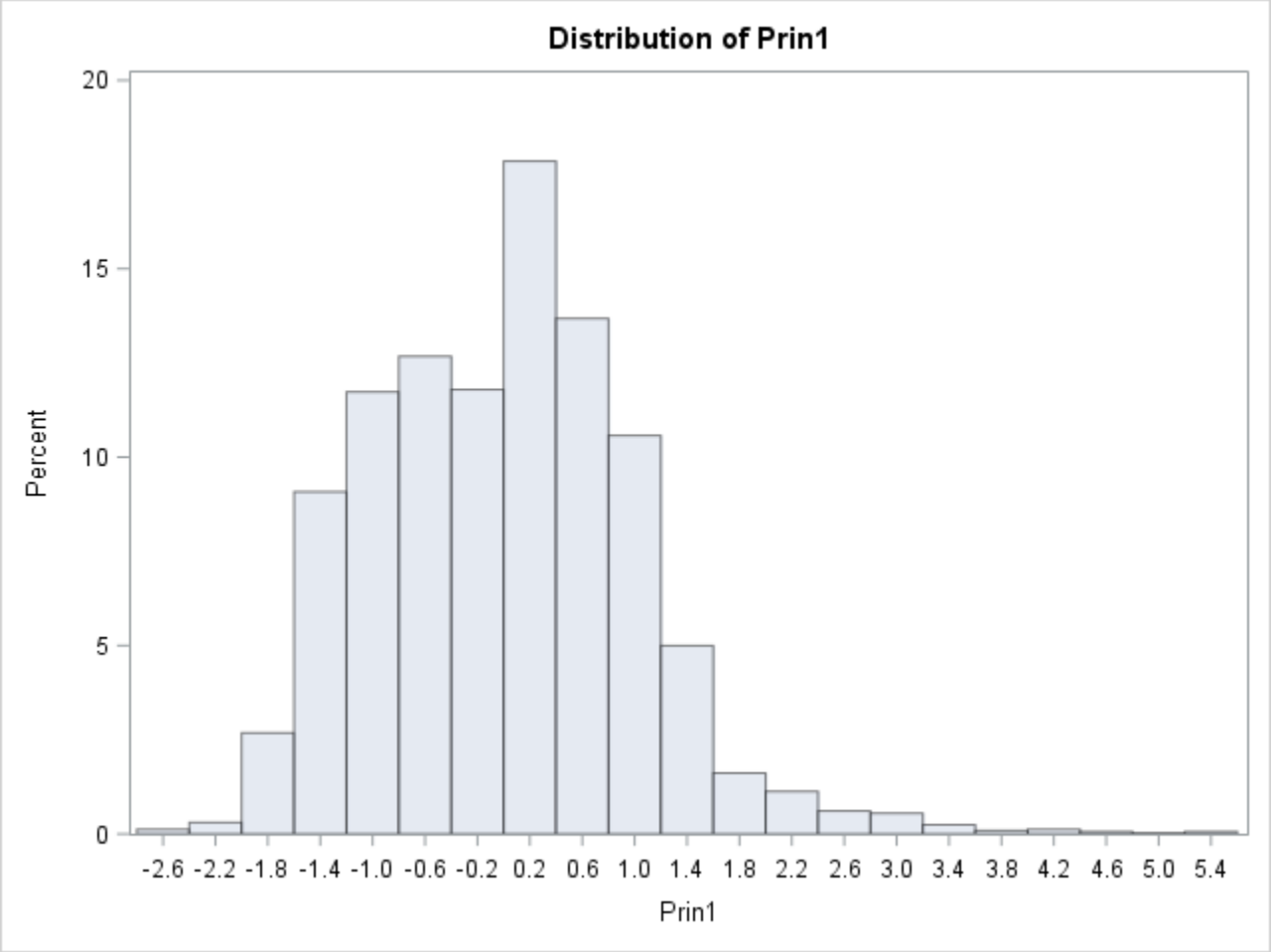


Table 8.9 *Post-Intervention Survey Principal Component Index 2 Defined Wealth Quintiles*

Index 2				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	657	20.02	657	20.02
1	650	19.80	1307	39.82
2	673	20.51	1980	60.33
3	646	19.68	2626	80.01
4	656	19.99	3282	100.00
Frequency Missing = 26				

Figure 8.6 *Post-Intervention Principal Component Index 3 Histogram*

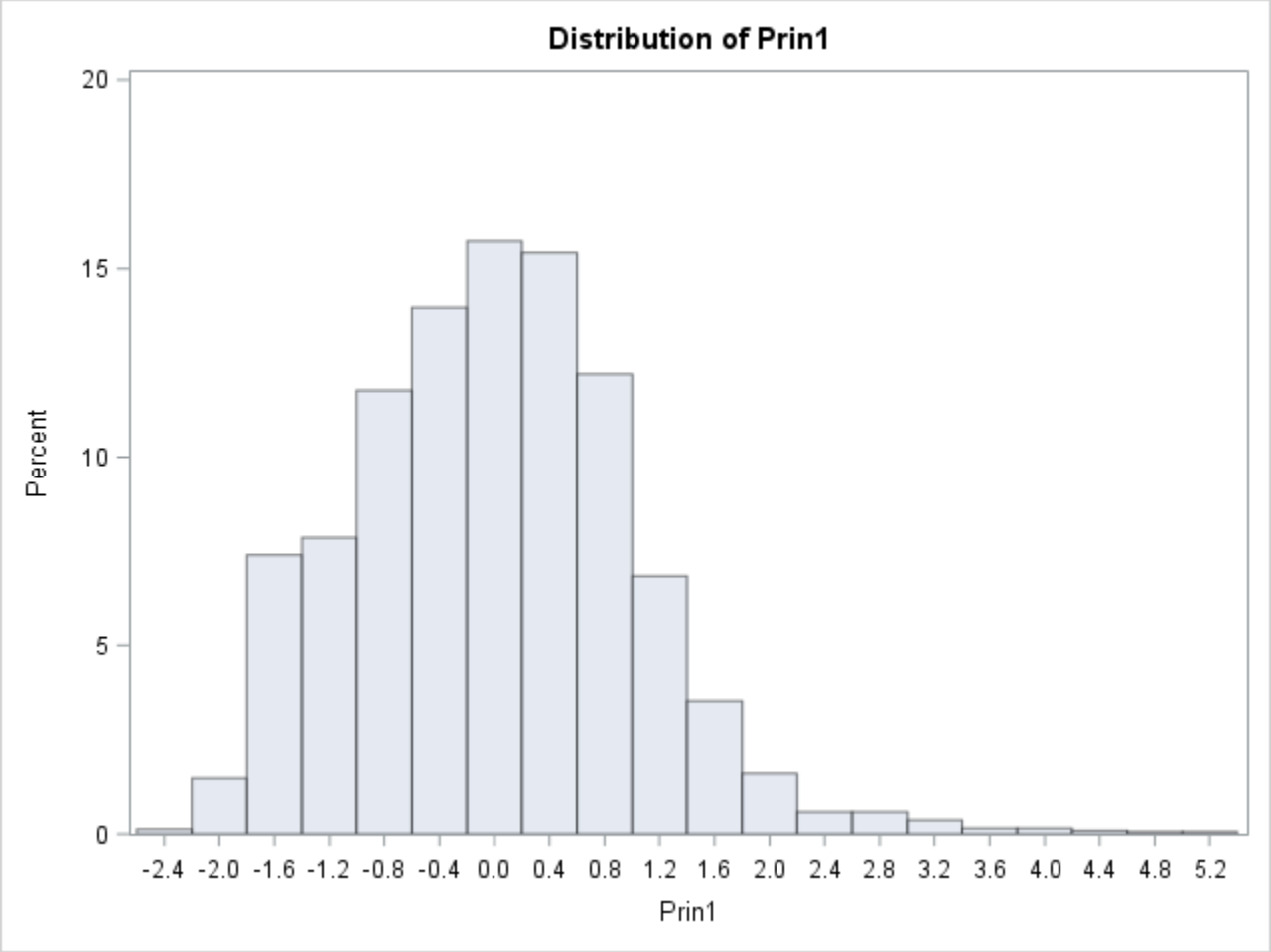


Table 8.10 *Post-Intervention Survey Principal Component Index 3 Defined Wealth Quintiles*

Index 3				
PCA	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	665	20.43	665	20.43
1	636	19.54	1301	39.97
2	653	20.06	1954	60.03
3	643	19.75	2597	79.78
4	658	20.22	3255	100.00
Frequency Missing = 53				

Figures 8.7 *Cluster 1 WLS Fitted Semivariogram for Residual Odds of Malaria*

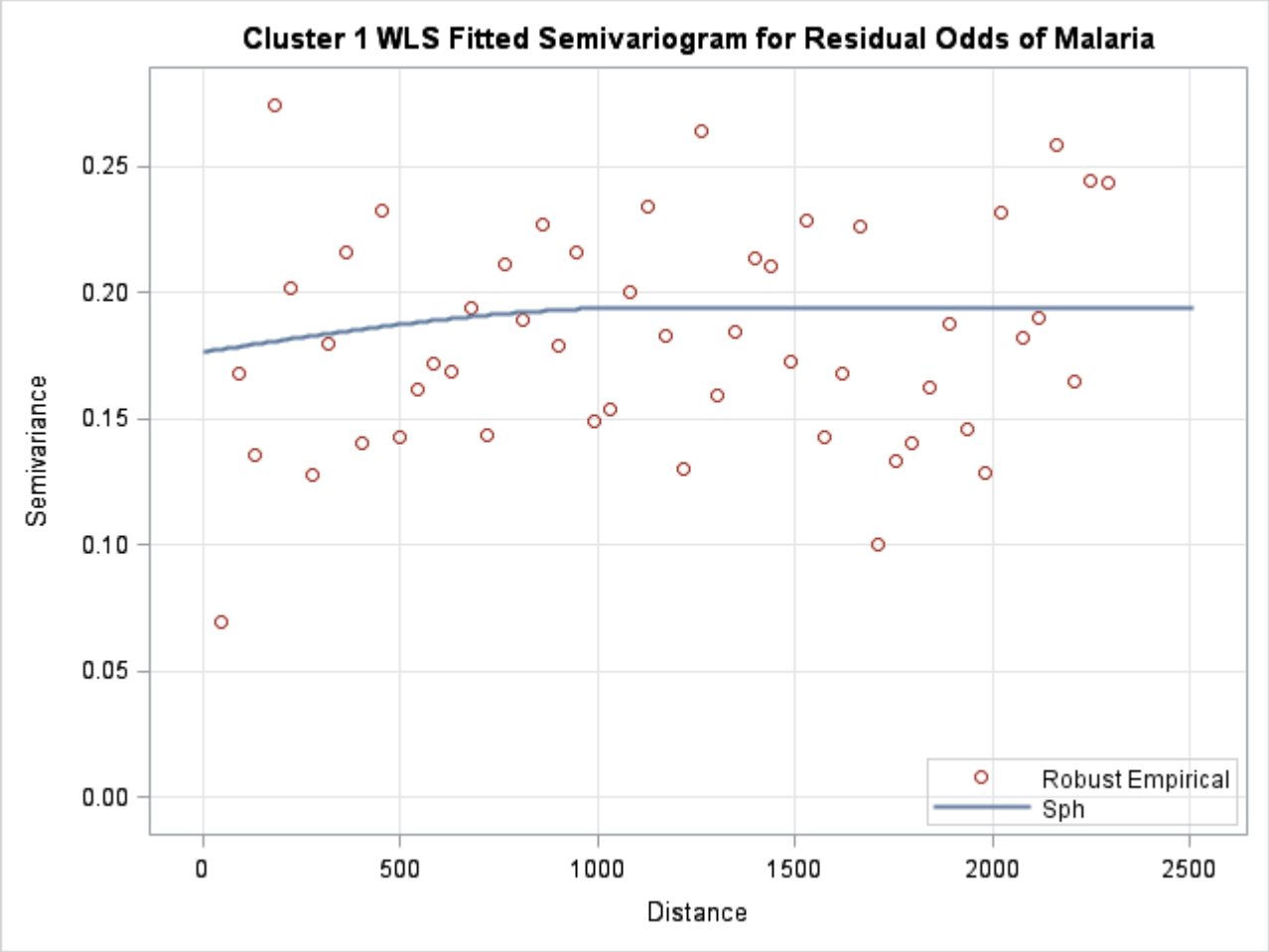


Figure 8.8 *Cluster 2 WLS Fitted Semivariogram for Residual Odds of Malaria*

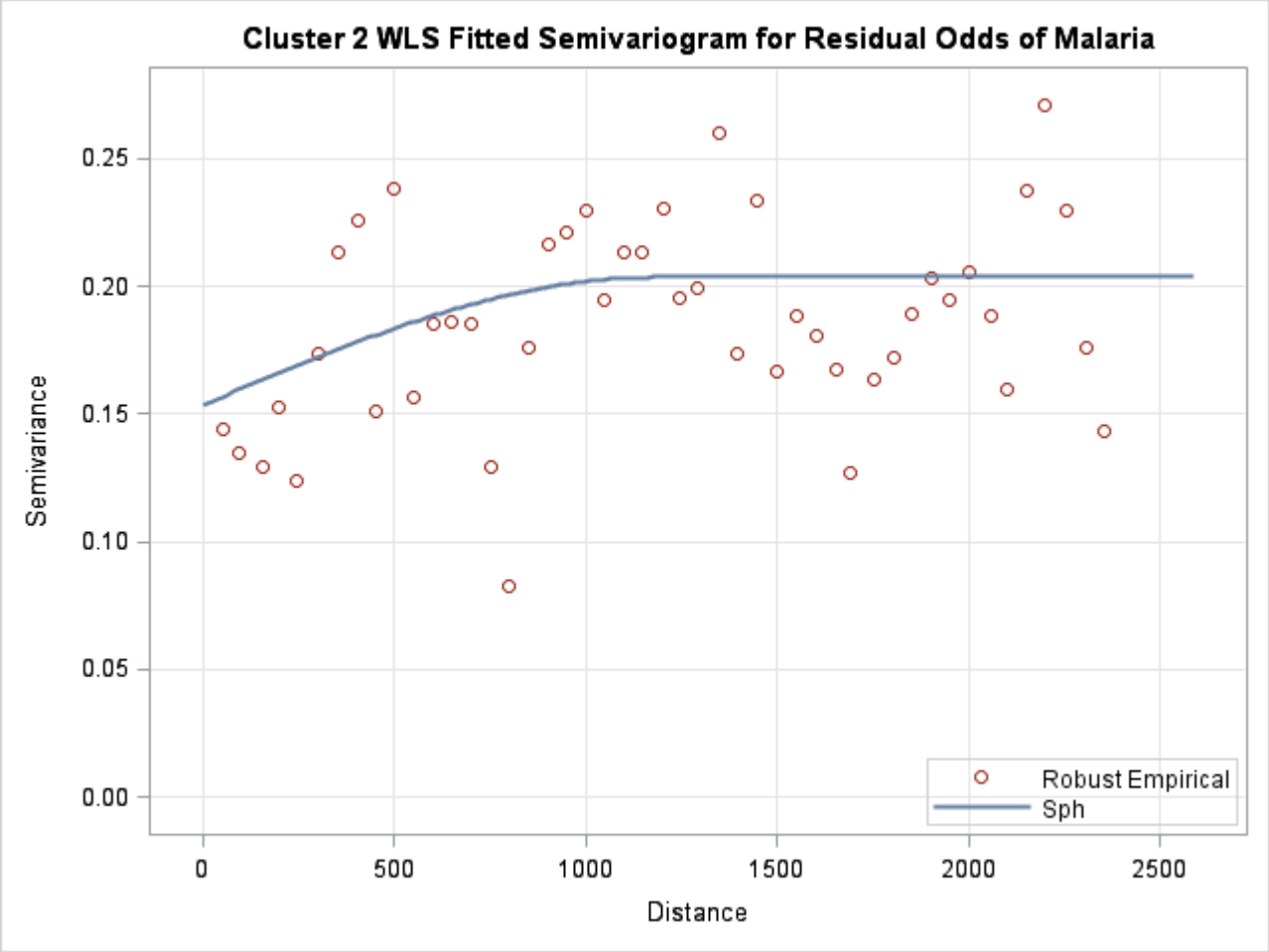


Figure 8.9 *Cluster 3 WLS Fitted Semivariogram for Residual Odds of Malaria*

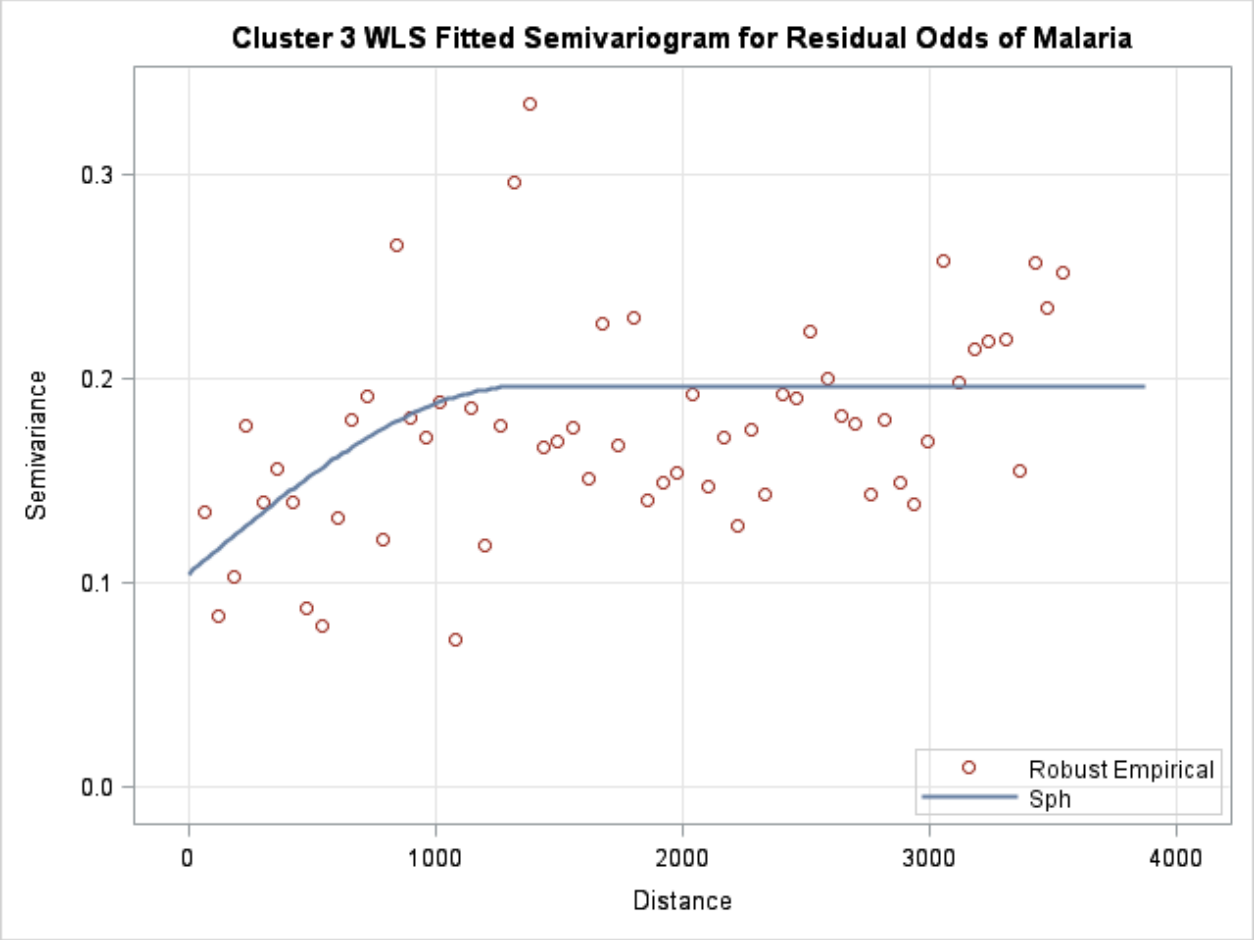


Figure 8.10 *Cluster 4 WLS Fitted Semivariogram for Residual Odds of Malaria*

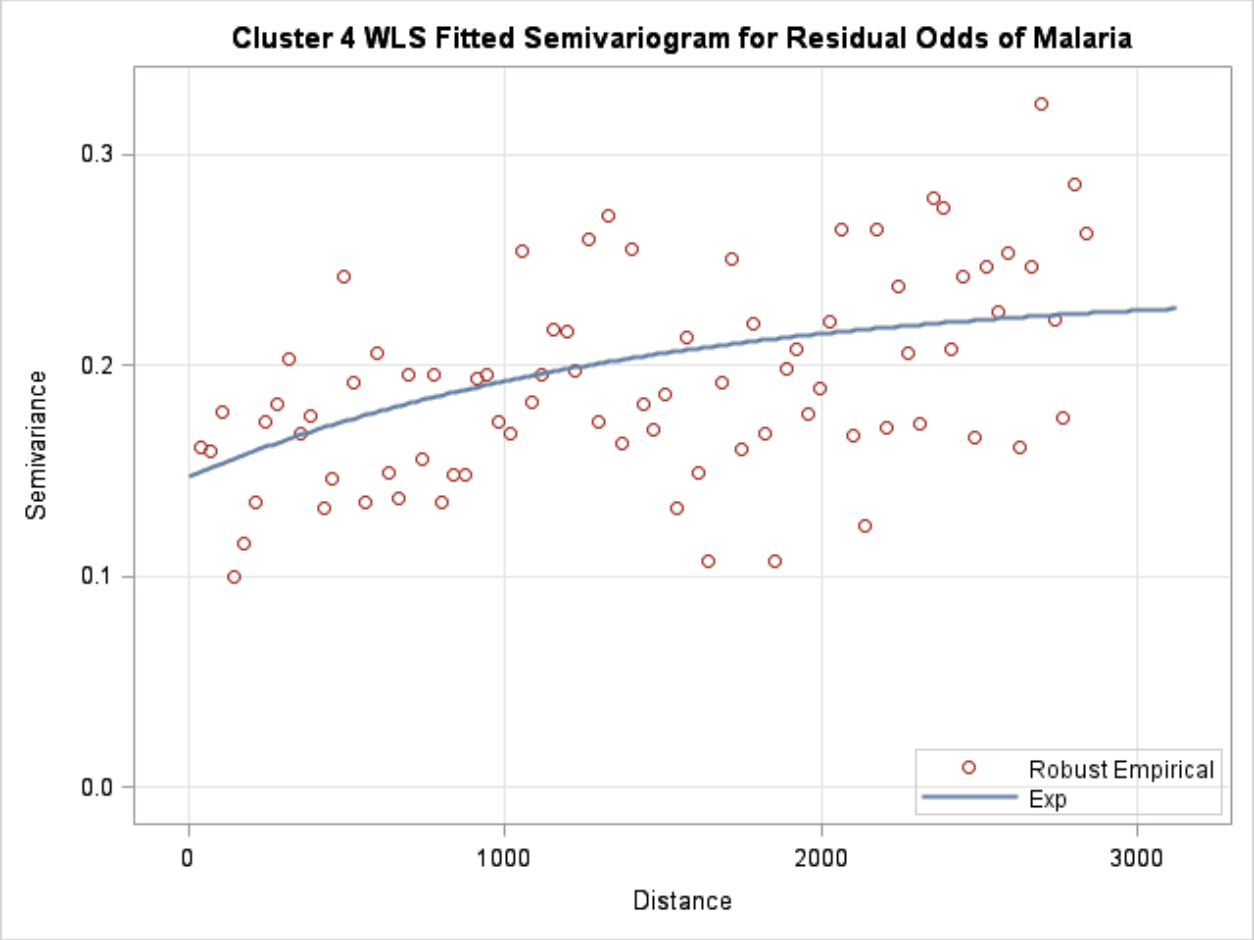


Figure 8.11 *Cluster 5 WLS Fitted Semivariogram for Residual Odds of Malaria*

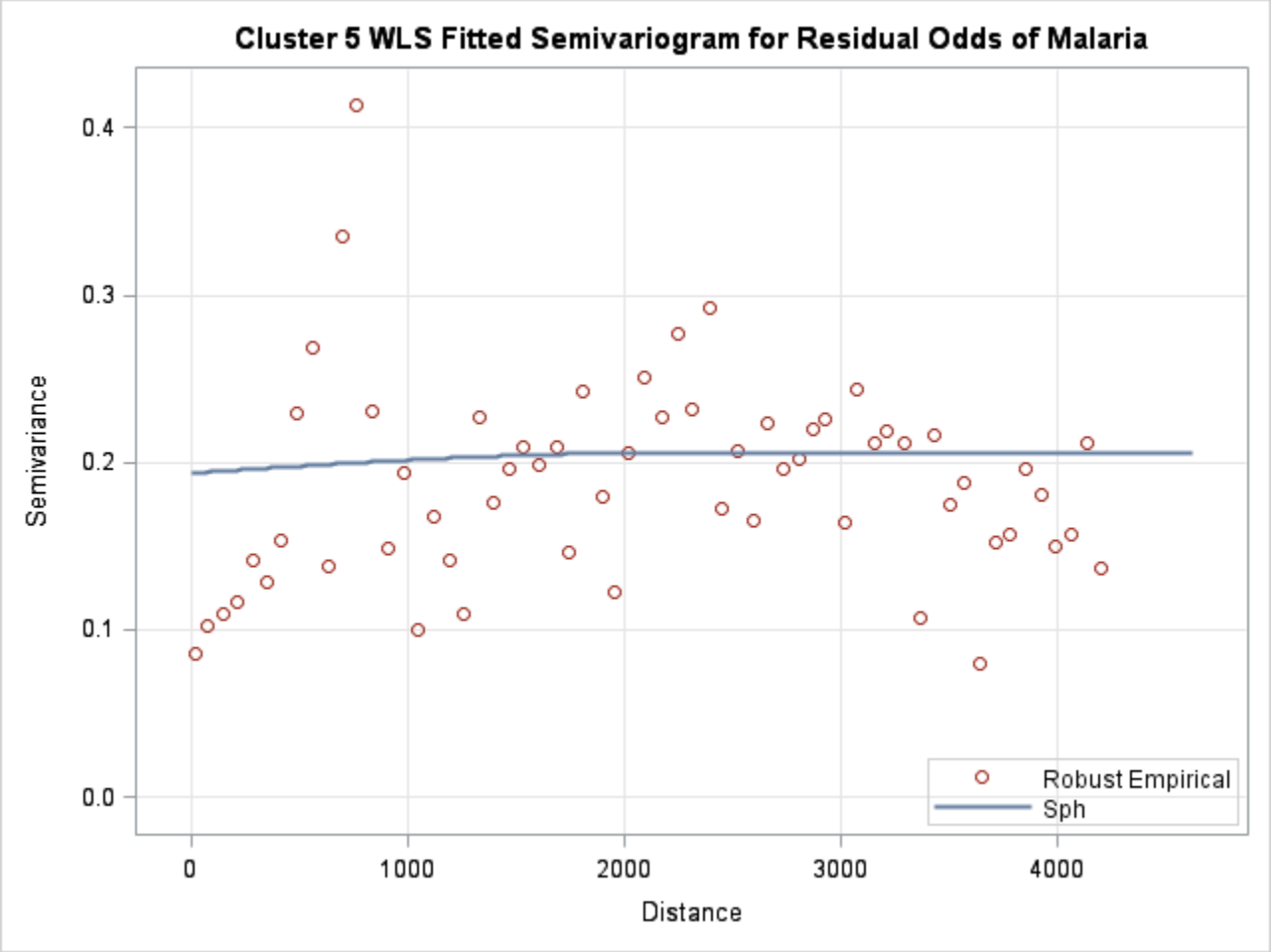


Figure 8.12 *Cluster 6 WLS Fitted Semivariogram for Residual Odds of Malaria*

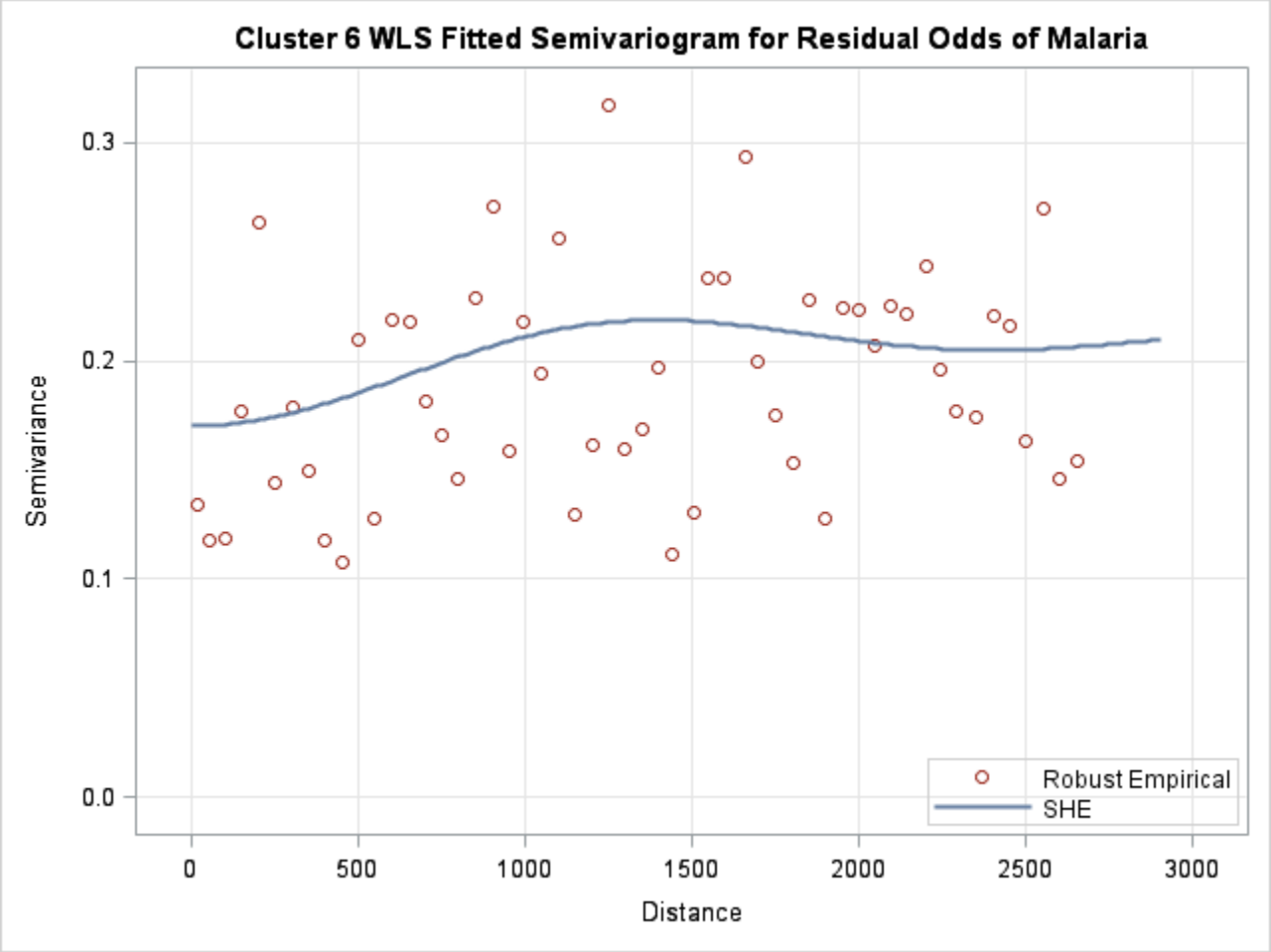


Figure 8.13 *Cluster 7 WLS Fitted Semivariogram for Residual Odds of Malaria*

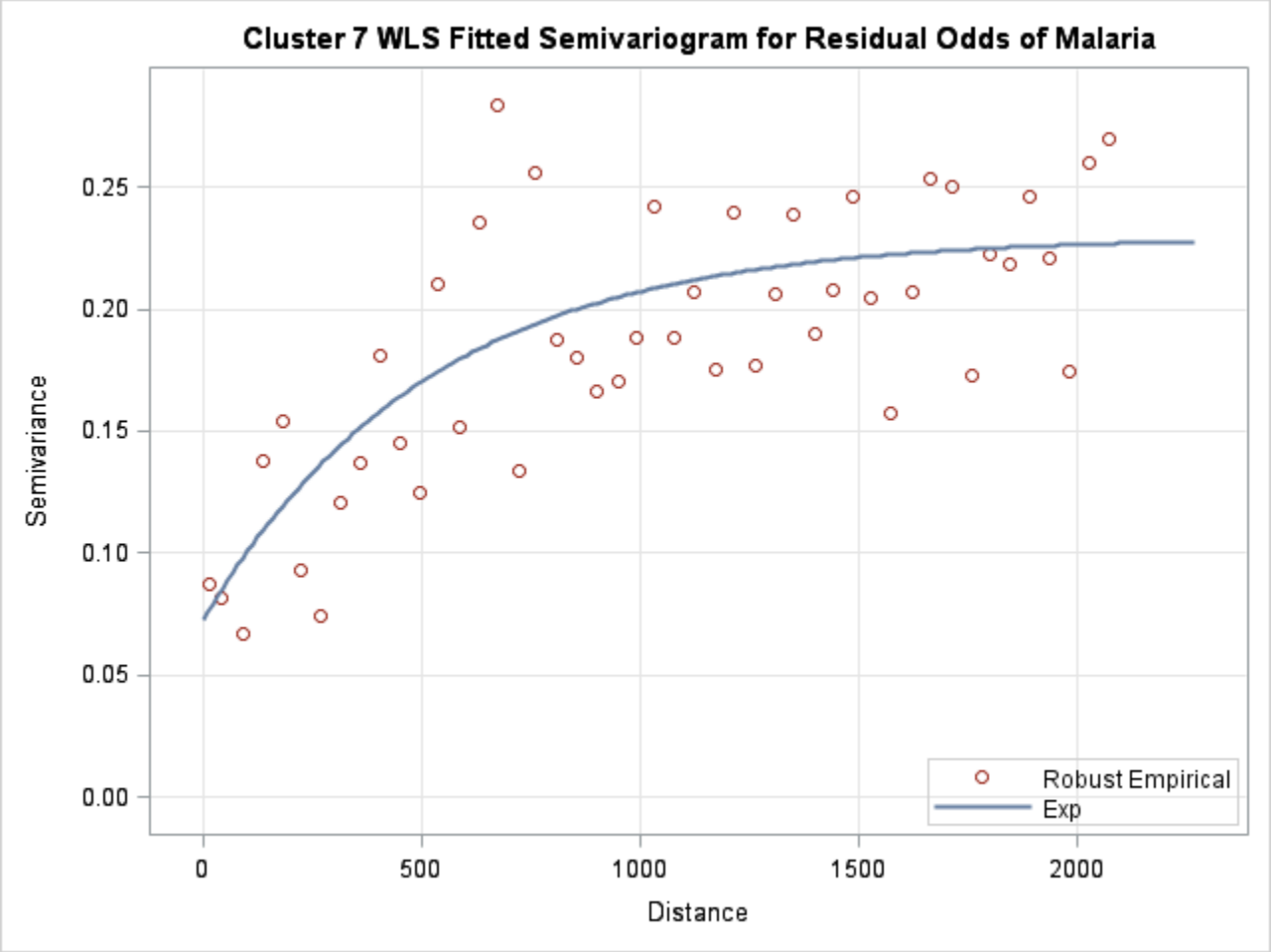


Figure 8.14 *Cluster 8 WLS Fitted Semivariogram for Residual Odds of Malaria*

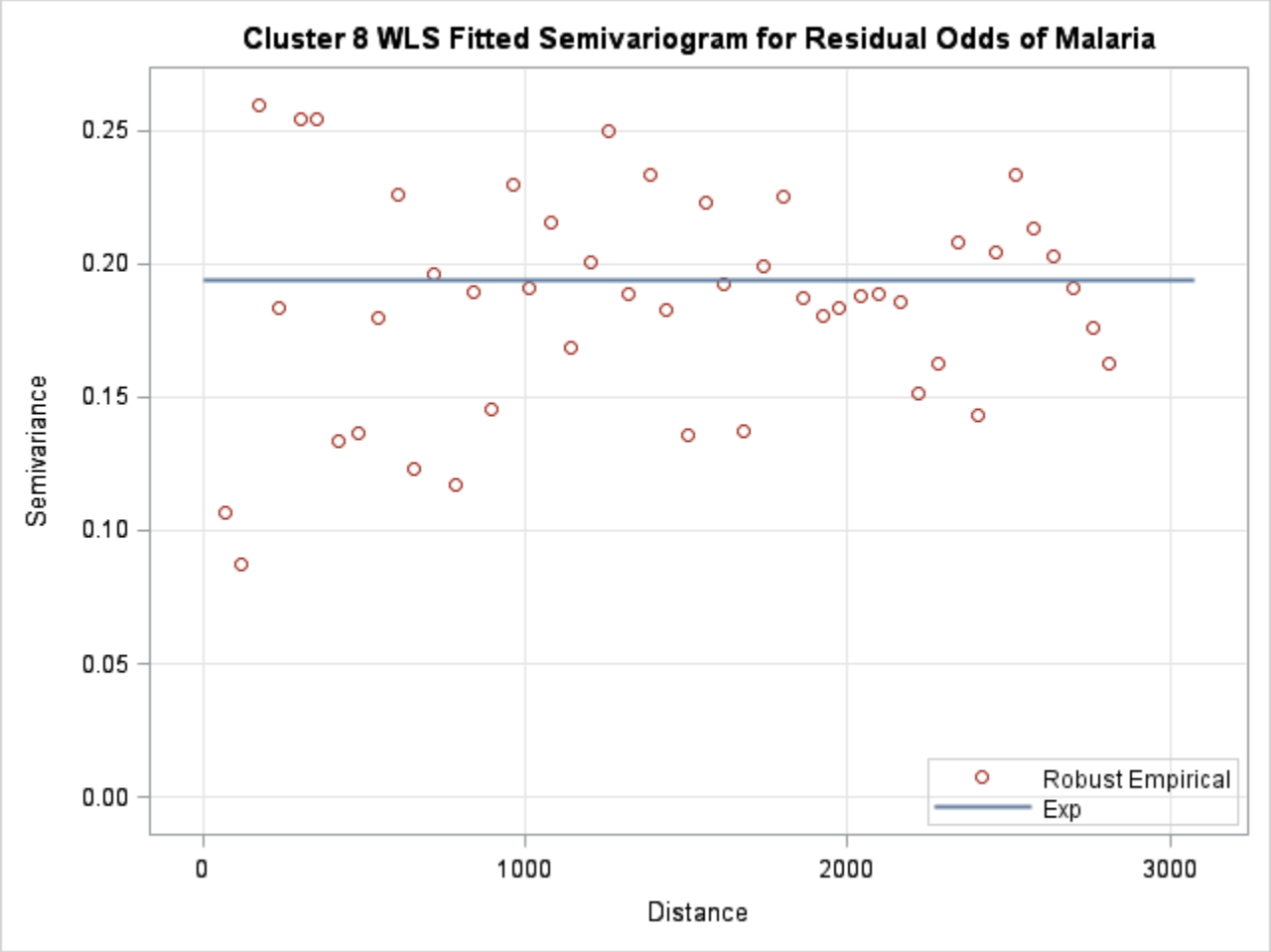


Figure 8.15 *Cluster 9 WLS Fitted Semivariogram for Residual Odds of Malaria*

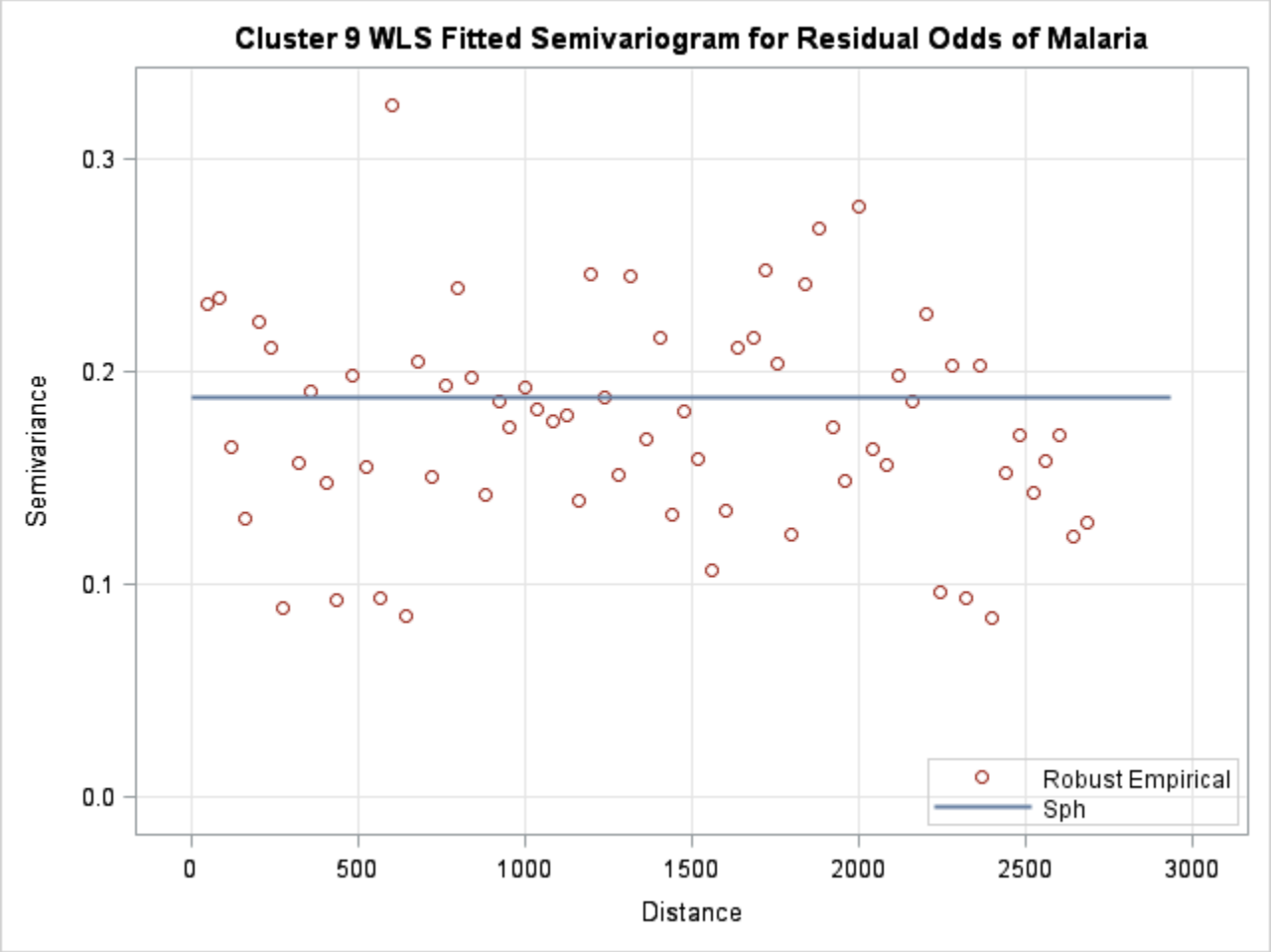


Figure 8.16 *Cluster 10 WLS Fitted Semivariogram for Residual Odds of Malaria*

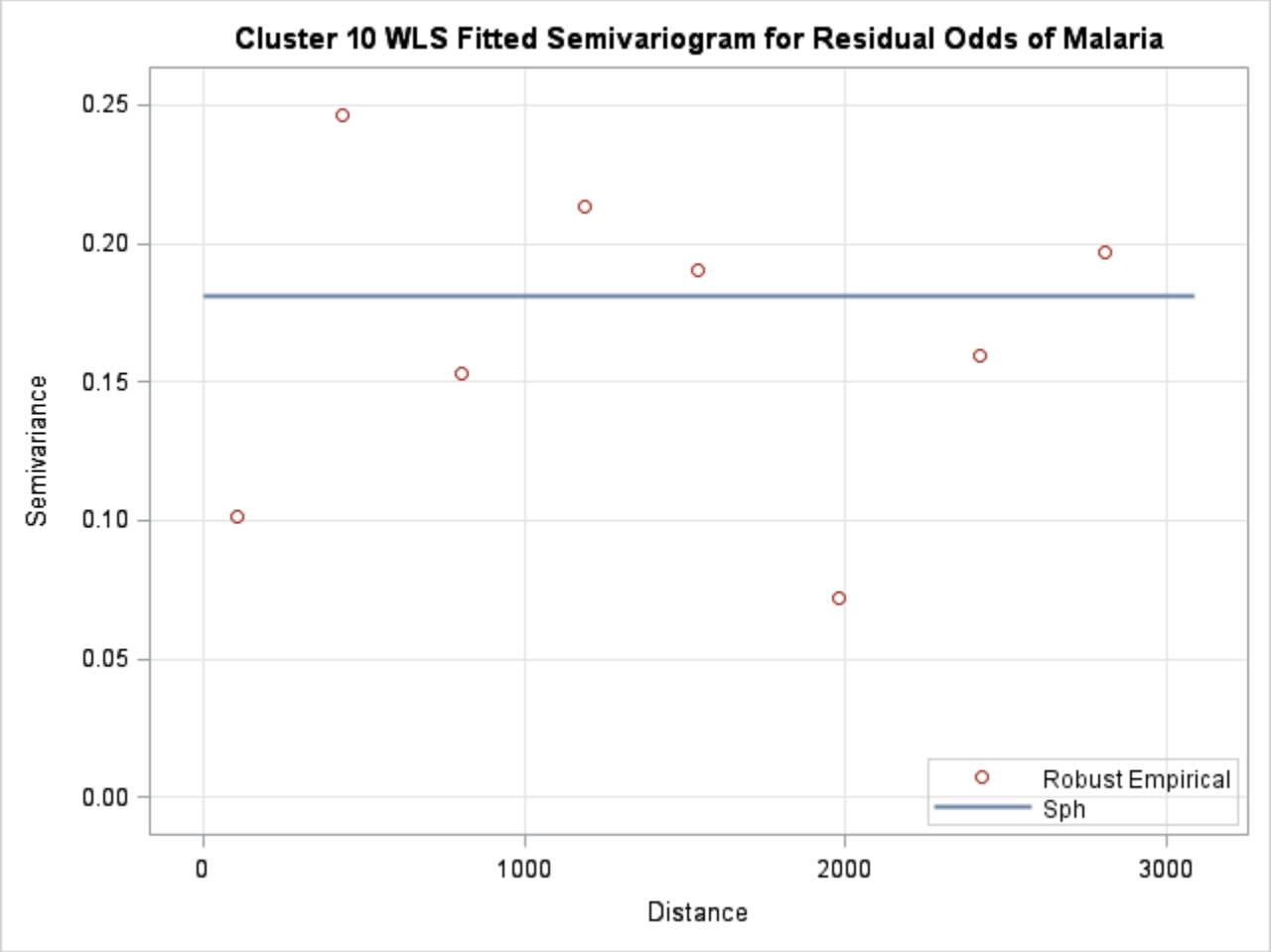


Figure 8.17 *Cluster 11 WLS Fitted Semivariogram for Residual Odds of Malaria*

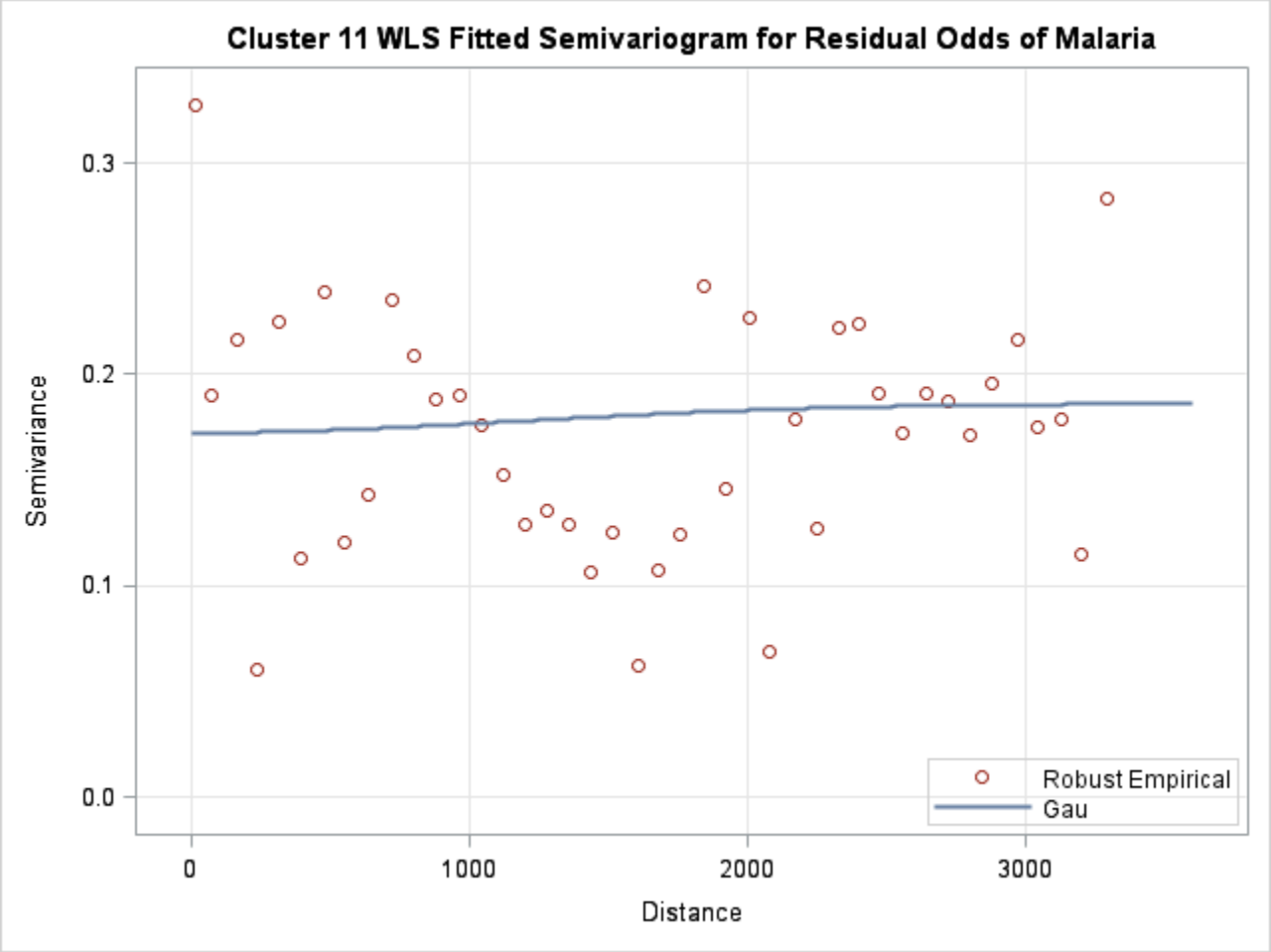


Figure 8.18 *Cluster 12 WLS Fitted Semivariogram for Residual Odds of Malaria*

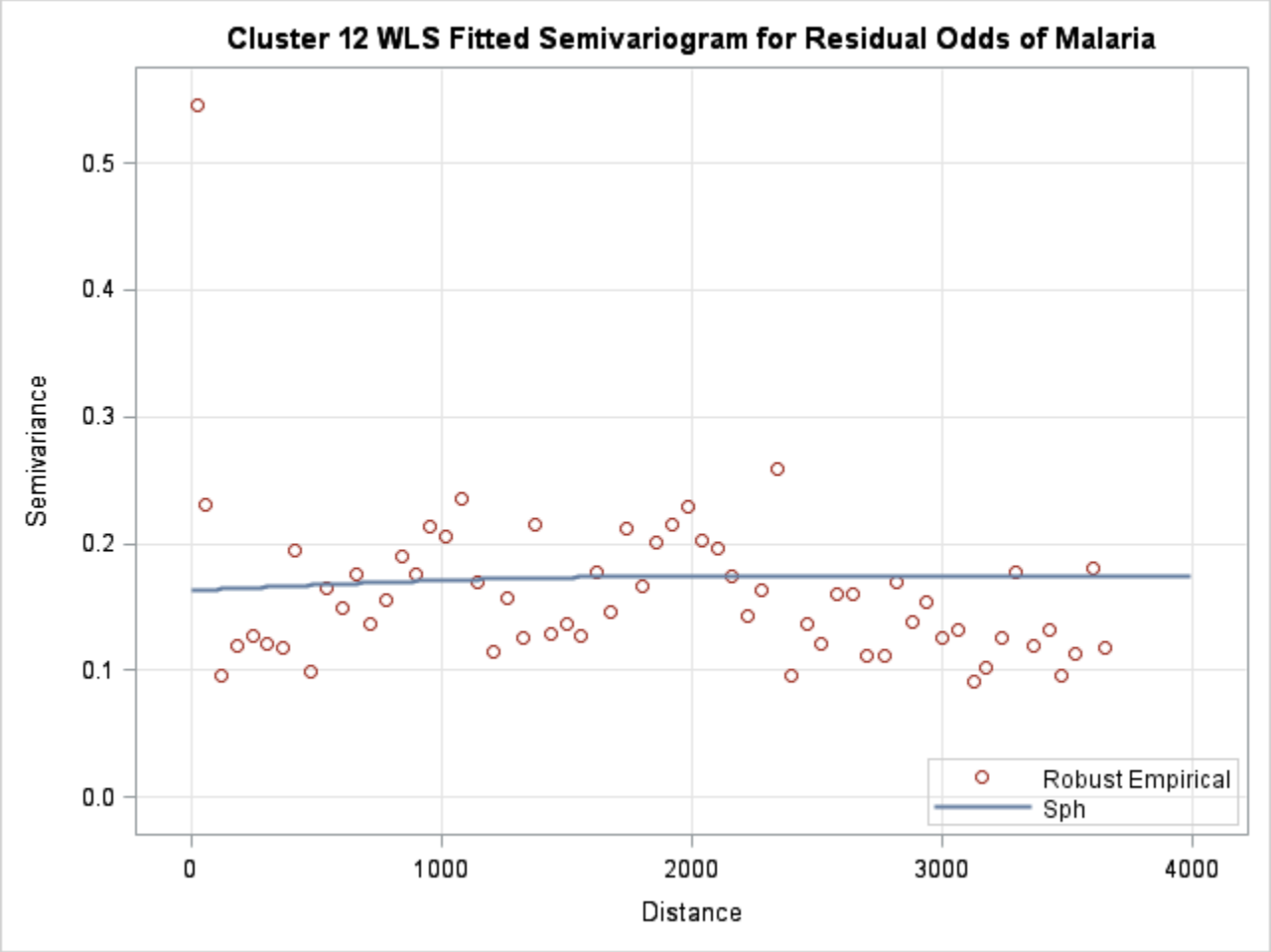


Figure 8.19 *Cluster 13 WLS Fitted Semivariogram for Residual Odds of Malaria*

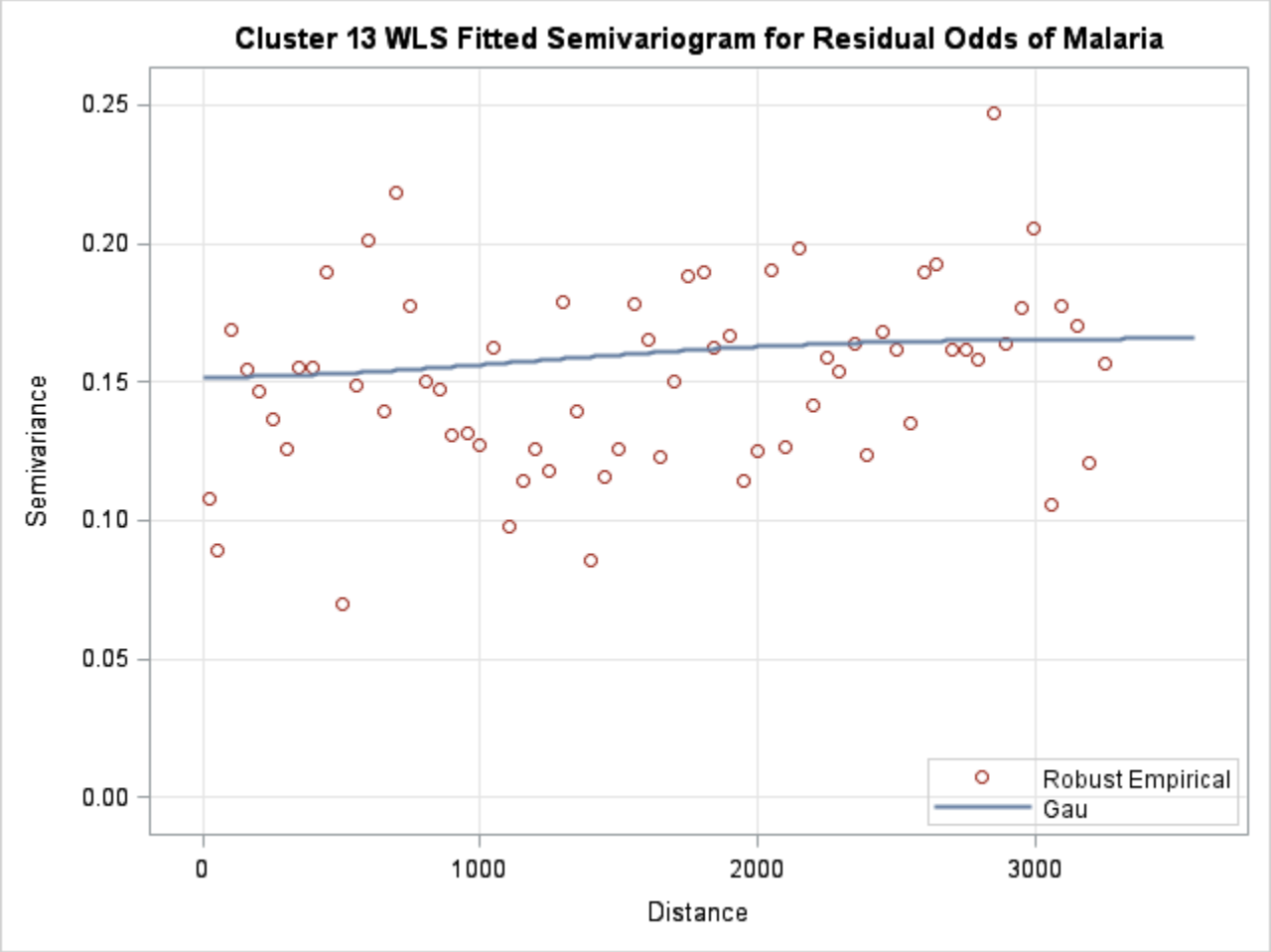


Figure 8.20 *Cluster 14 WLS Fitted Semivariogram for Residual Odds of Malaria*

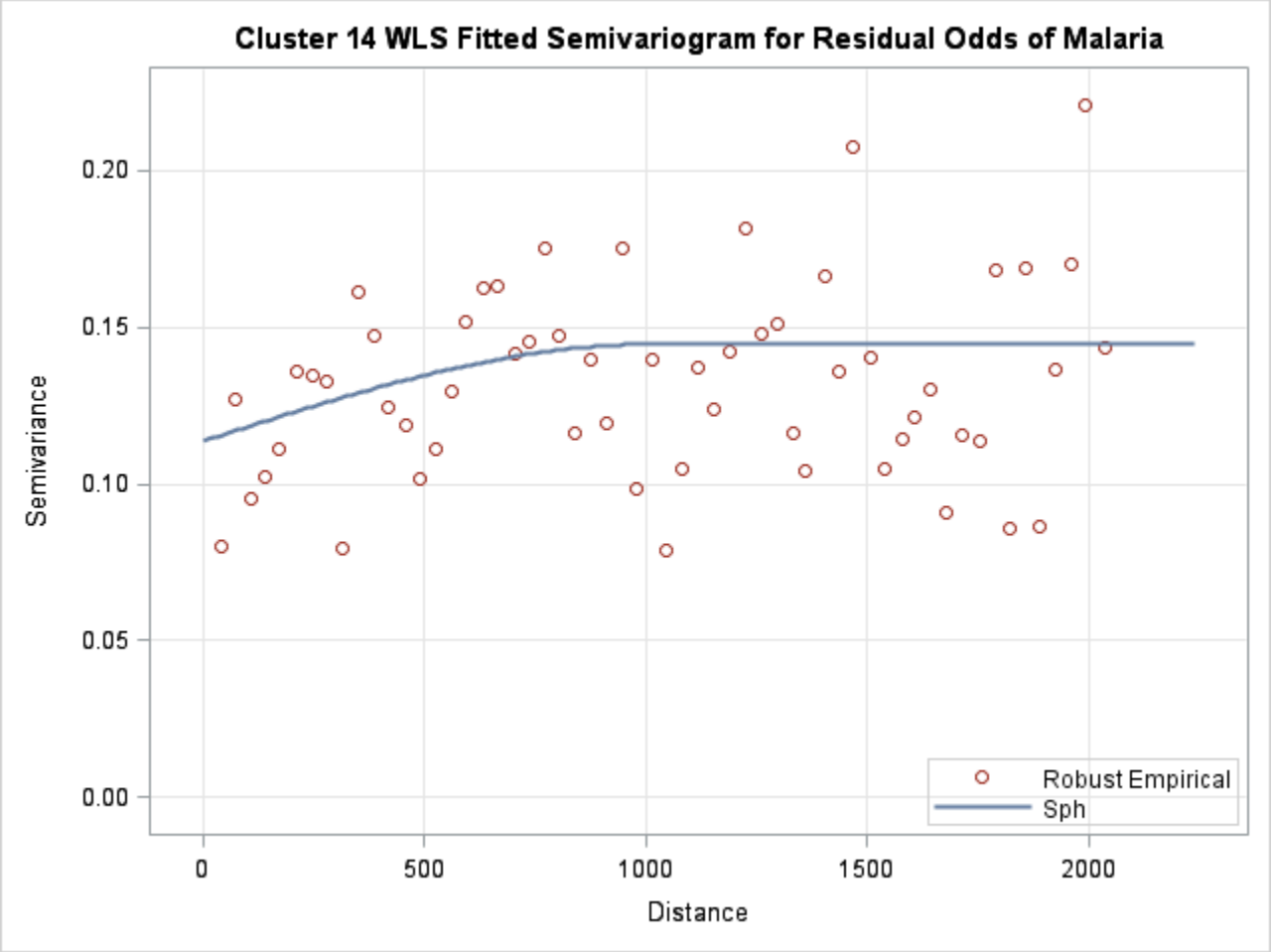


Figure 8.21 *Cluster 15 WLS Fitted Semivariogram for Residual Odds of Malaria*

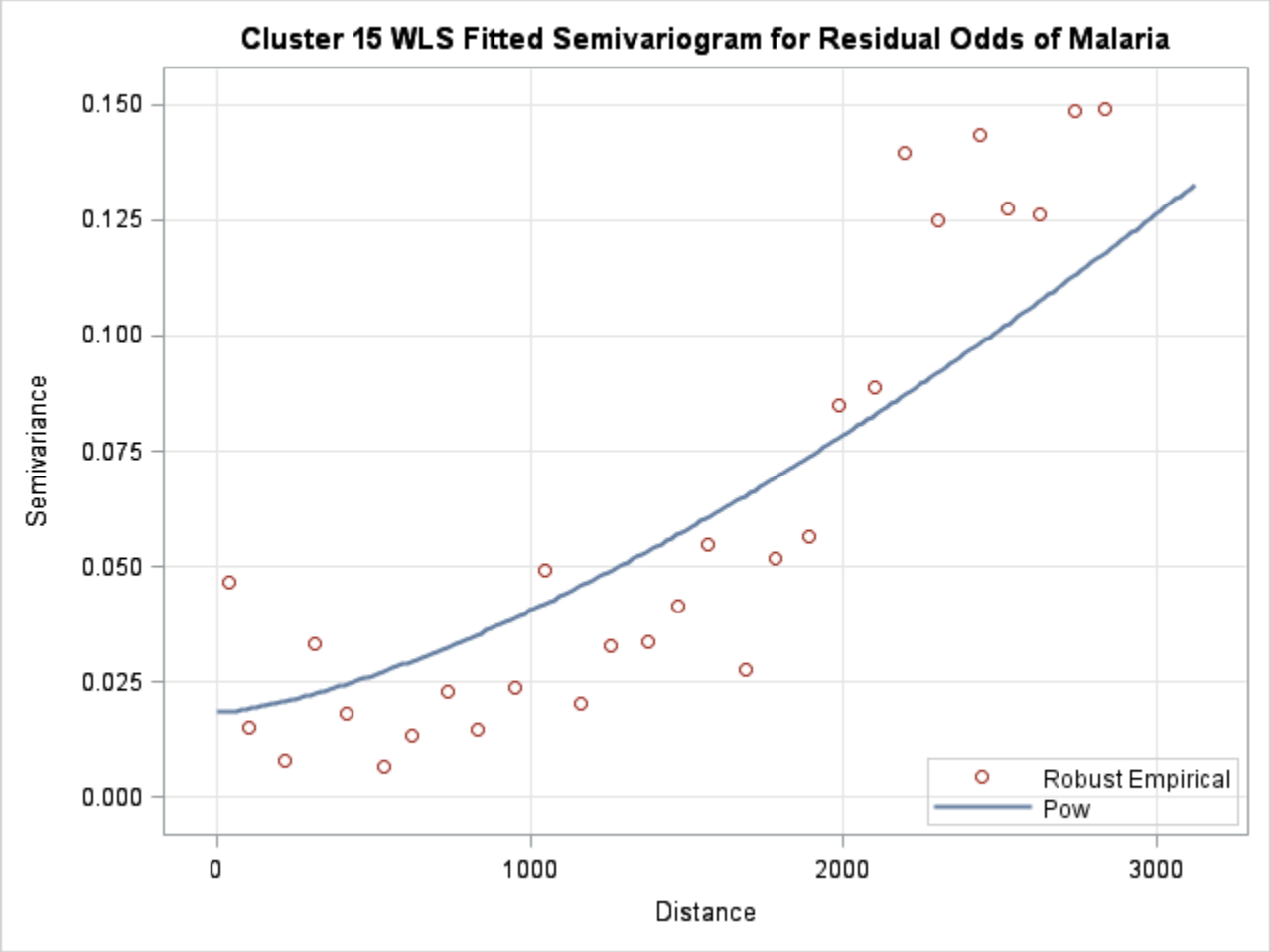


Figure 8.22 *Cluster 16 WLS Fitted Semivariogram for Residual Odds of Malaria*

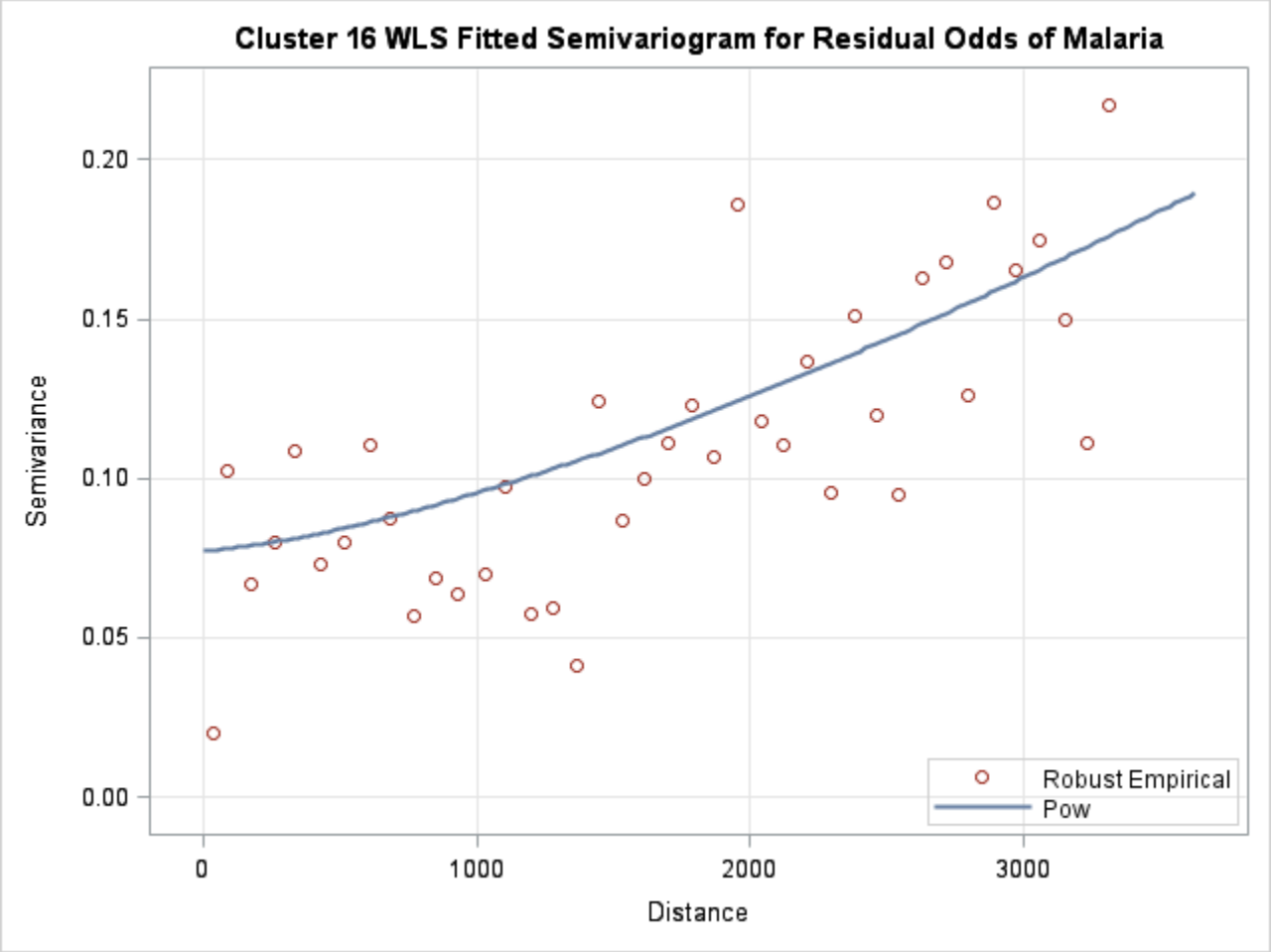


Figure 8.23 *Cluster 17 WLS Fitted Semivariogram for Residual Odds of Malaria*

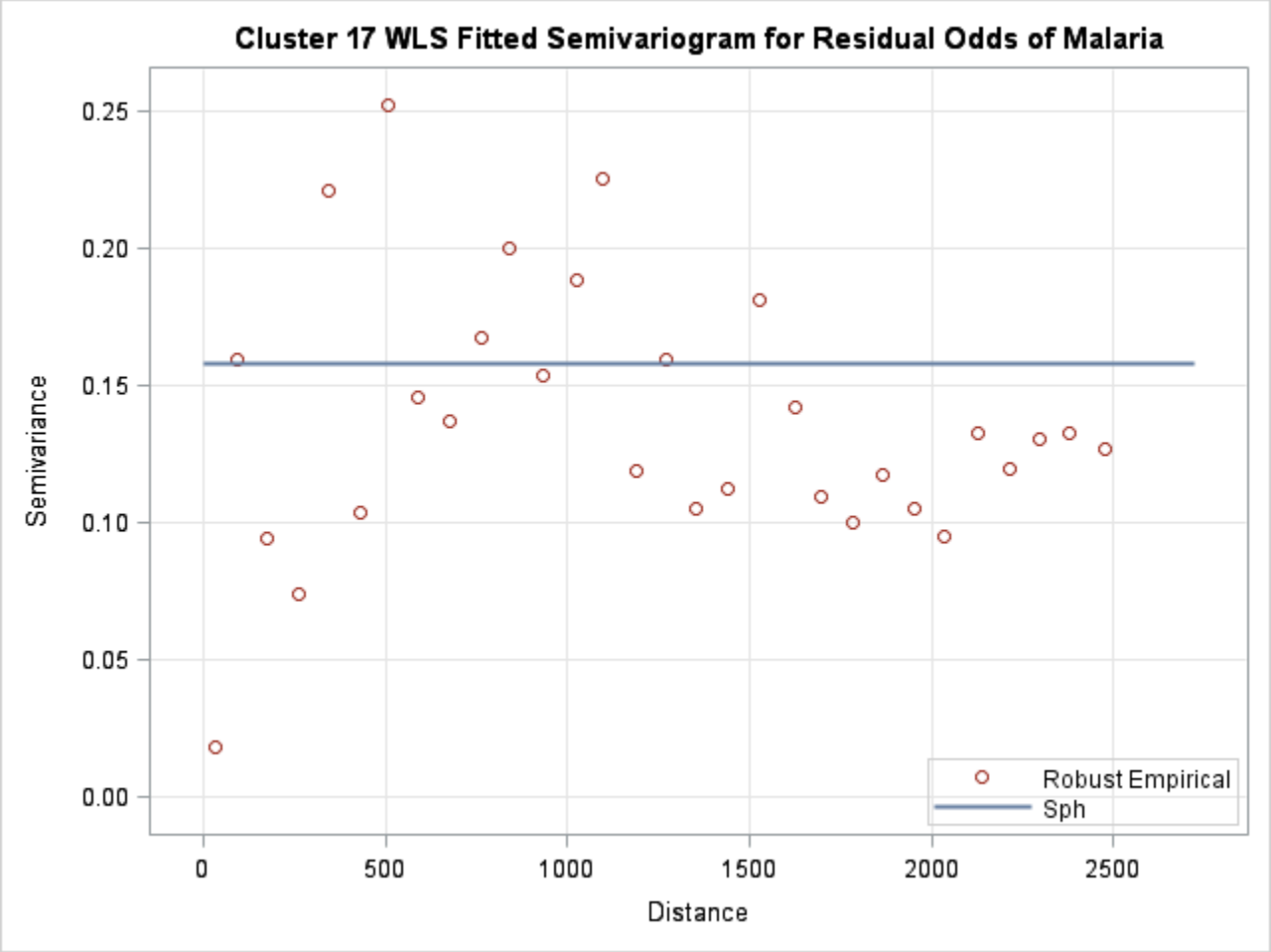


Figure 8.24 *Cluster 18 WLS Fitted Semivariogram for Residual Odds of Malaria*

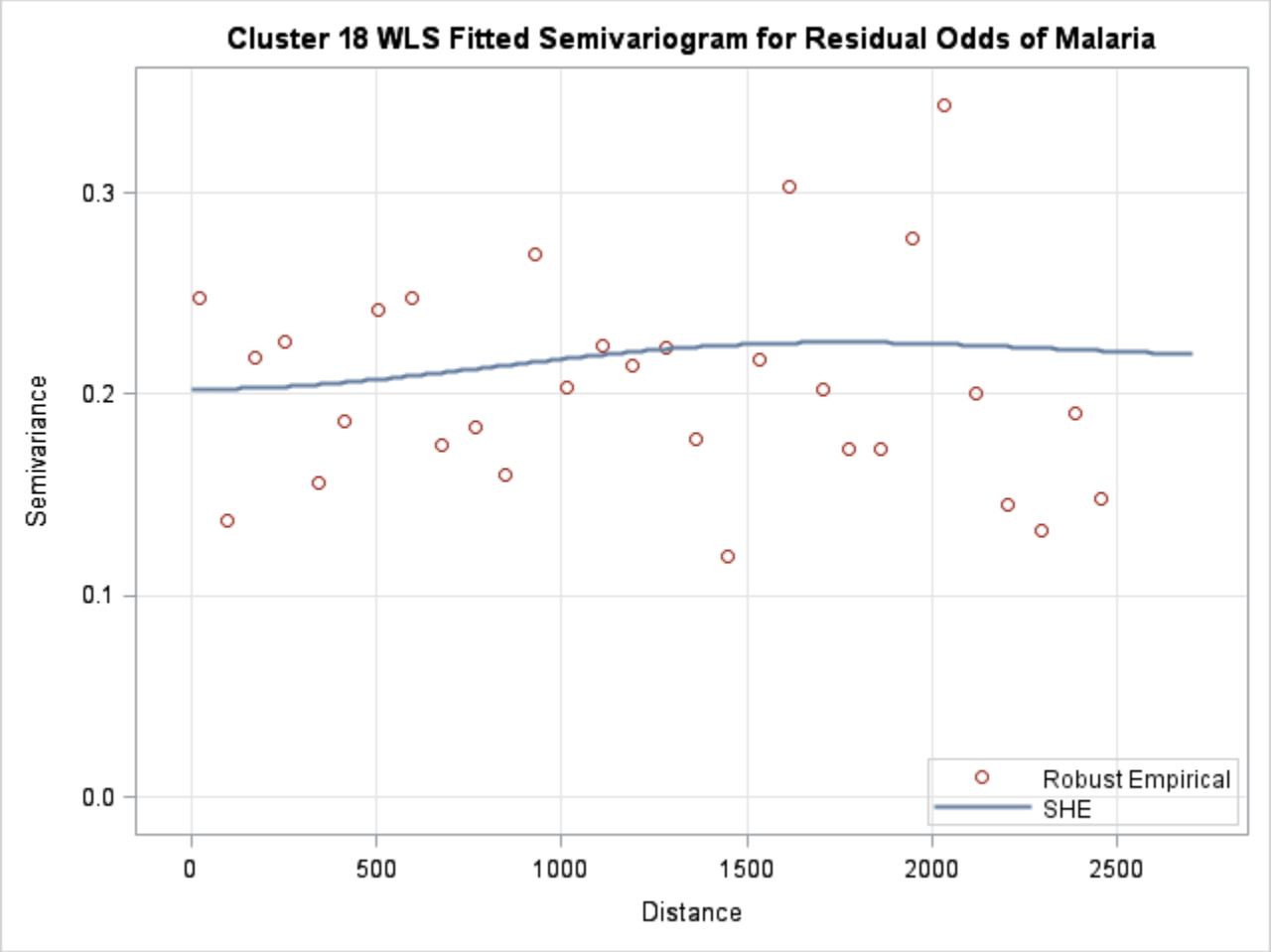


Figure 8.25 *Cluster 19 WLS Fitted Semivariogram for Residual Odds of Malaria*

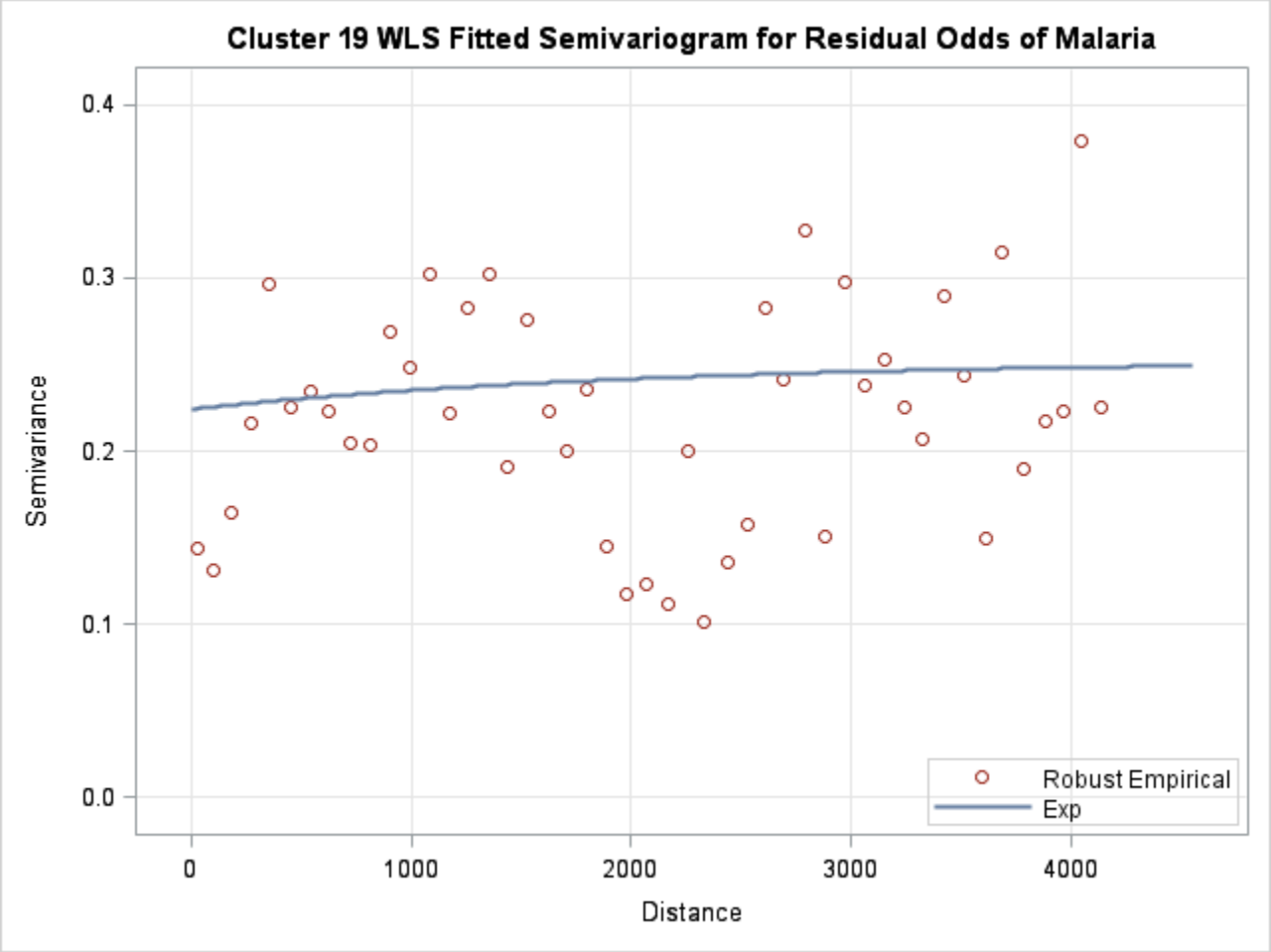


Figure 8.26 *Cluster 20 WLS Fitted Semivariogram for Residual Odds of Malaria*

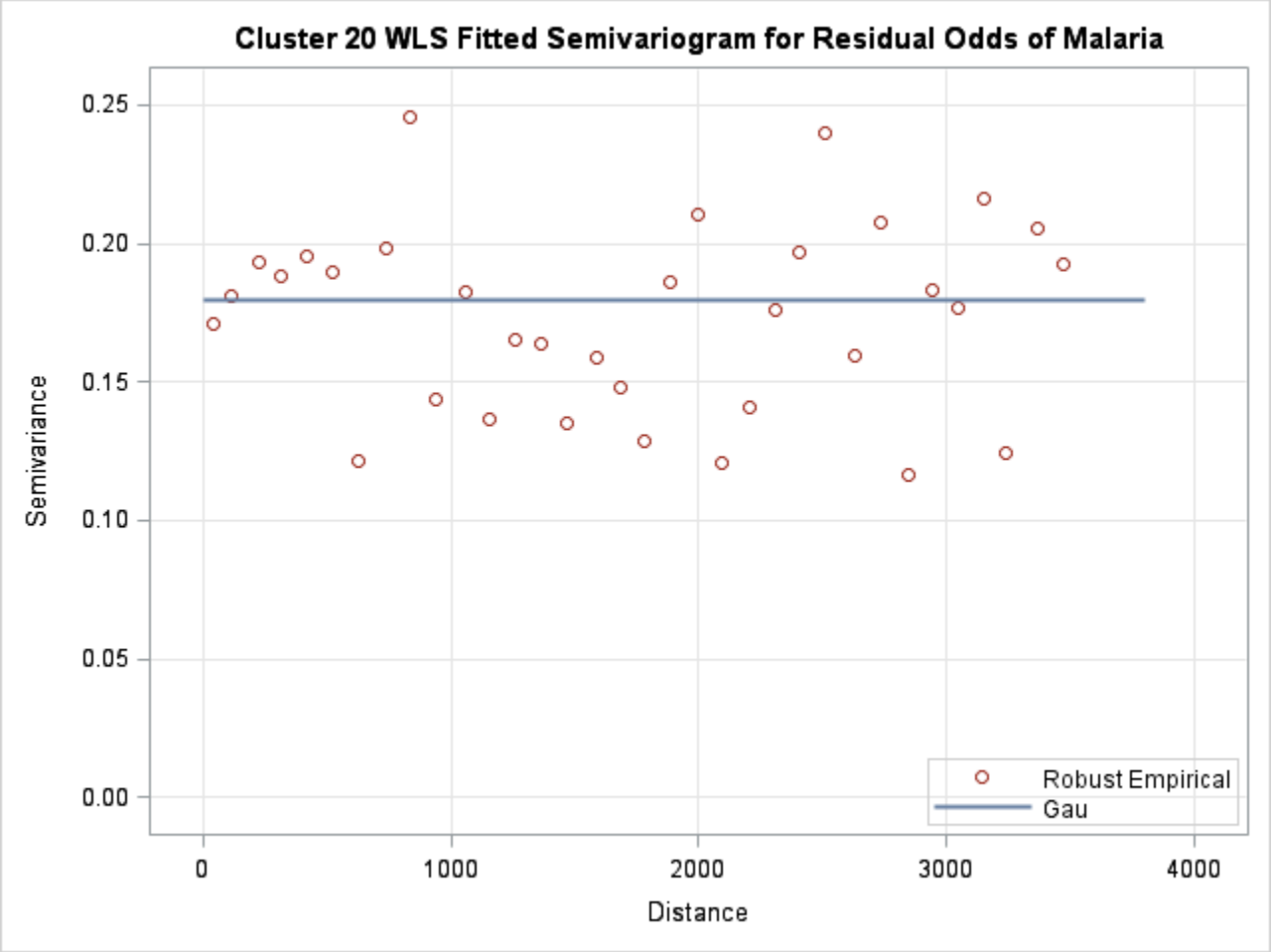


Figure 8.27 *Cluster 21 WLS Fitted Semivariogram for Residual Odds of Malaria*

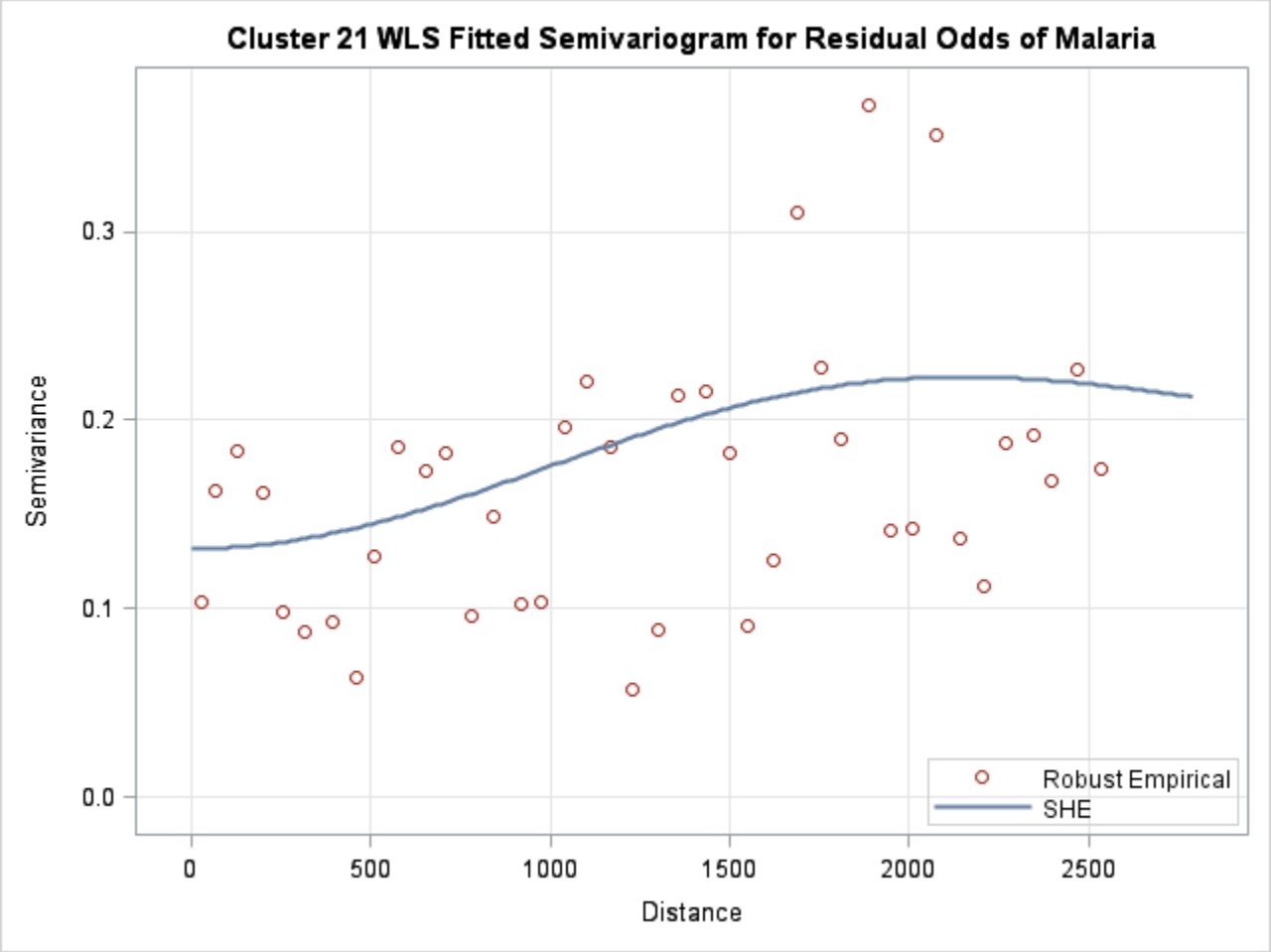


Figure 8.28 *Cluster 22 WLS Fitted Semivariogram for Residual Odds of Malaria*

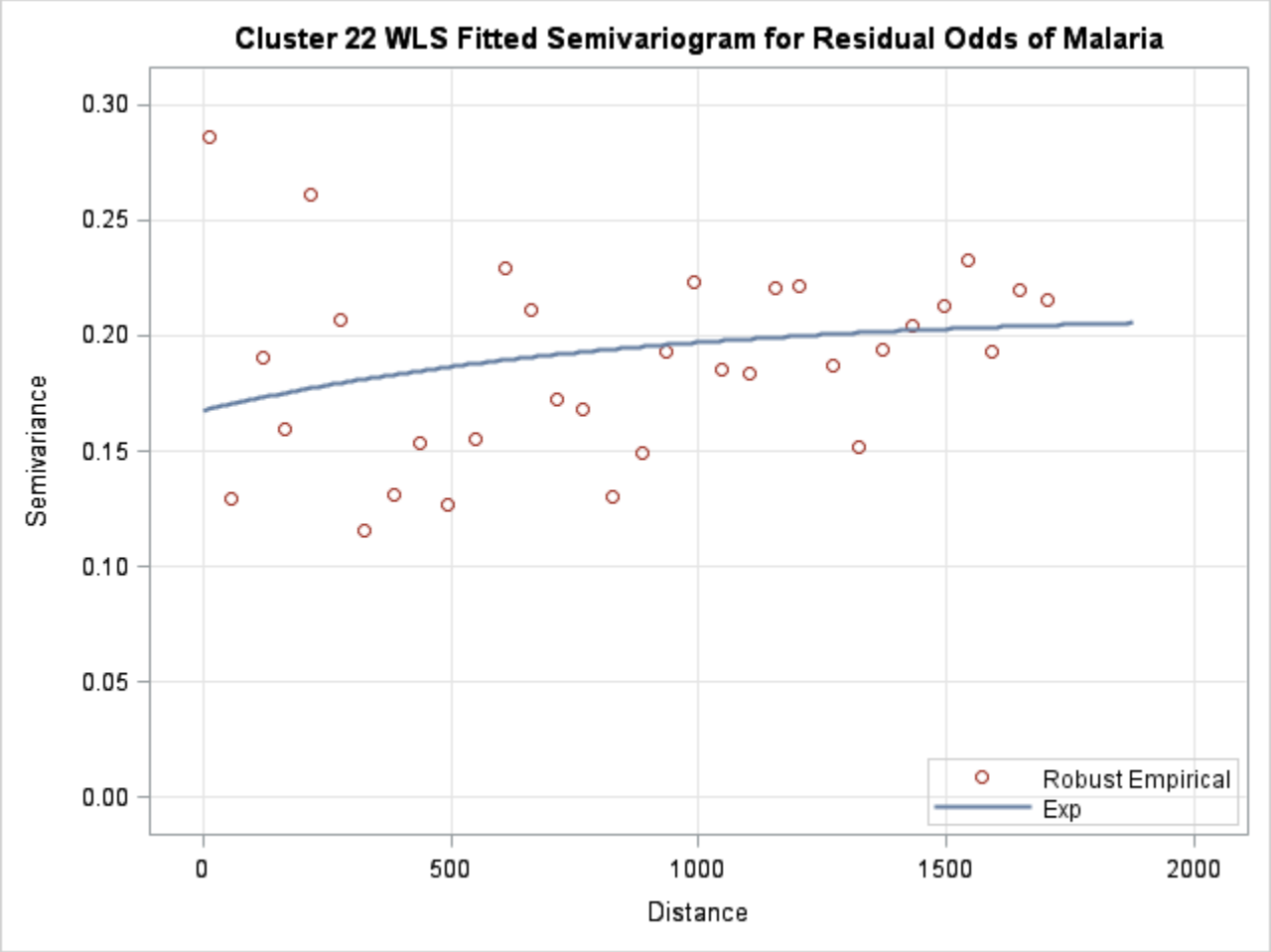


Figure 8.29 *Cluster 23 WLS Fitted Semivariogram for Residual Odds of Malaria*

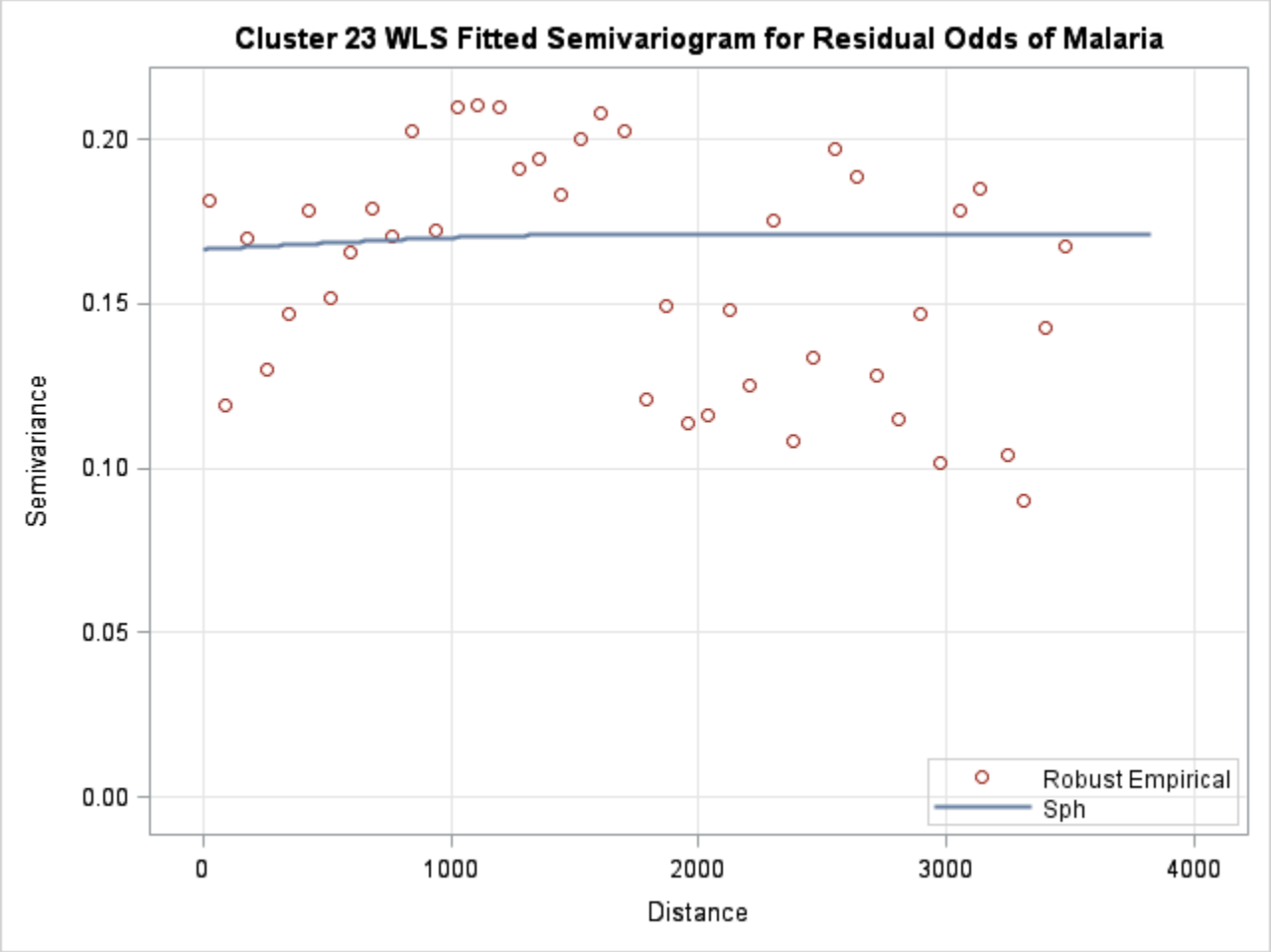


Figure 8.30 *Cluster 24 WLS Fitted Semivariogram for Residual Odds of Malaria*

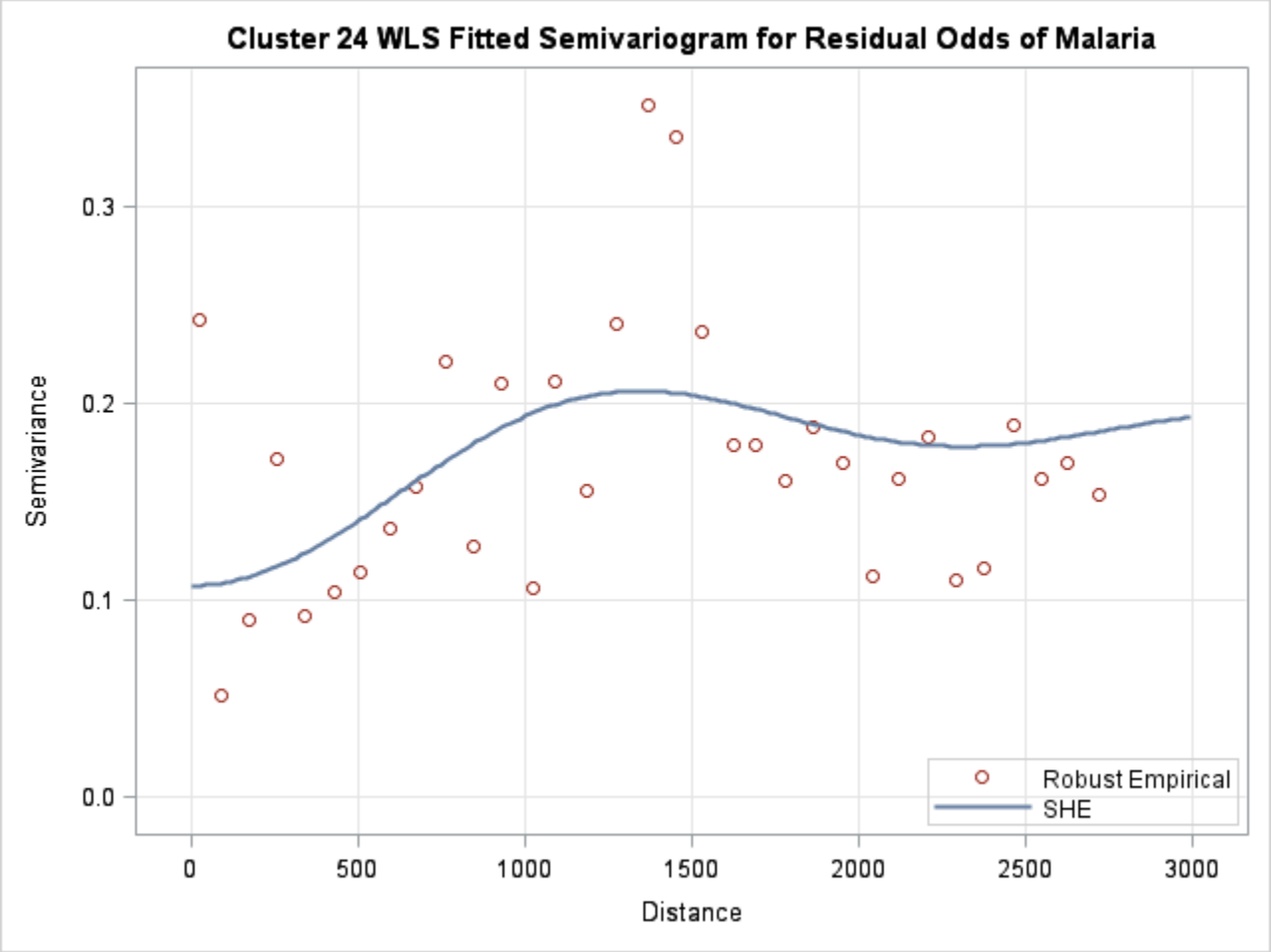


Figure 8.31 *Cluster 25 WLS Fitted Semivariogram for Residual Odds of Malaria*

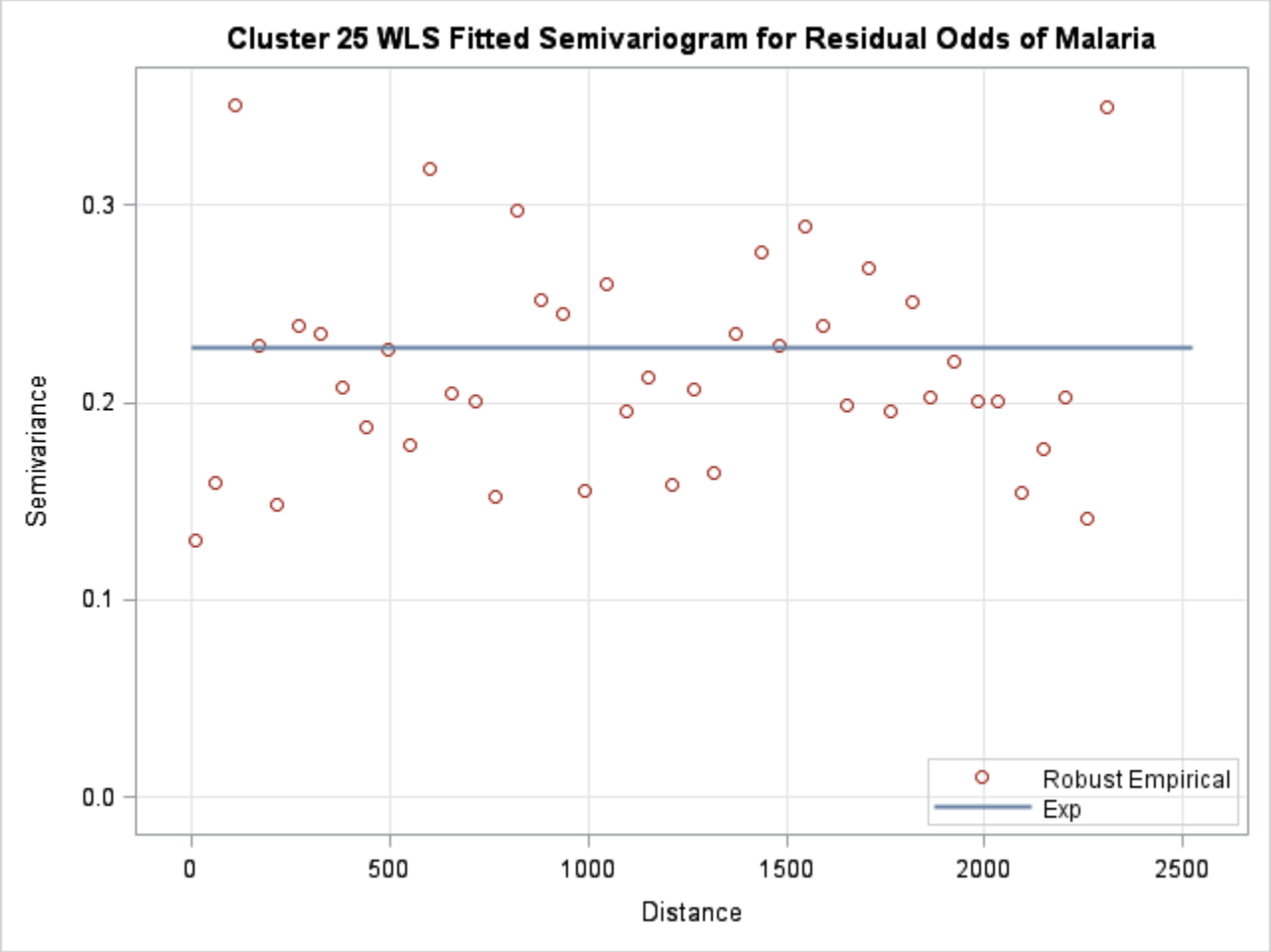


Figure 8.32 *Cluster 26 WLS Fitted Semivariogram for Residual Odds of Malaria*

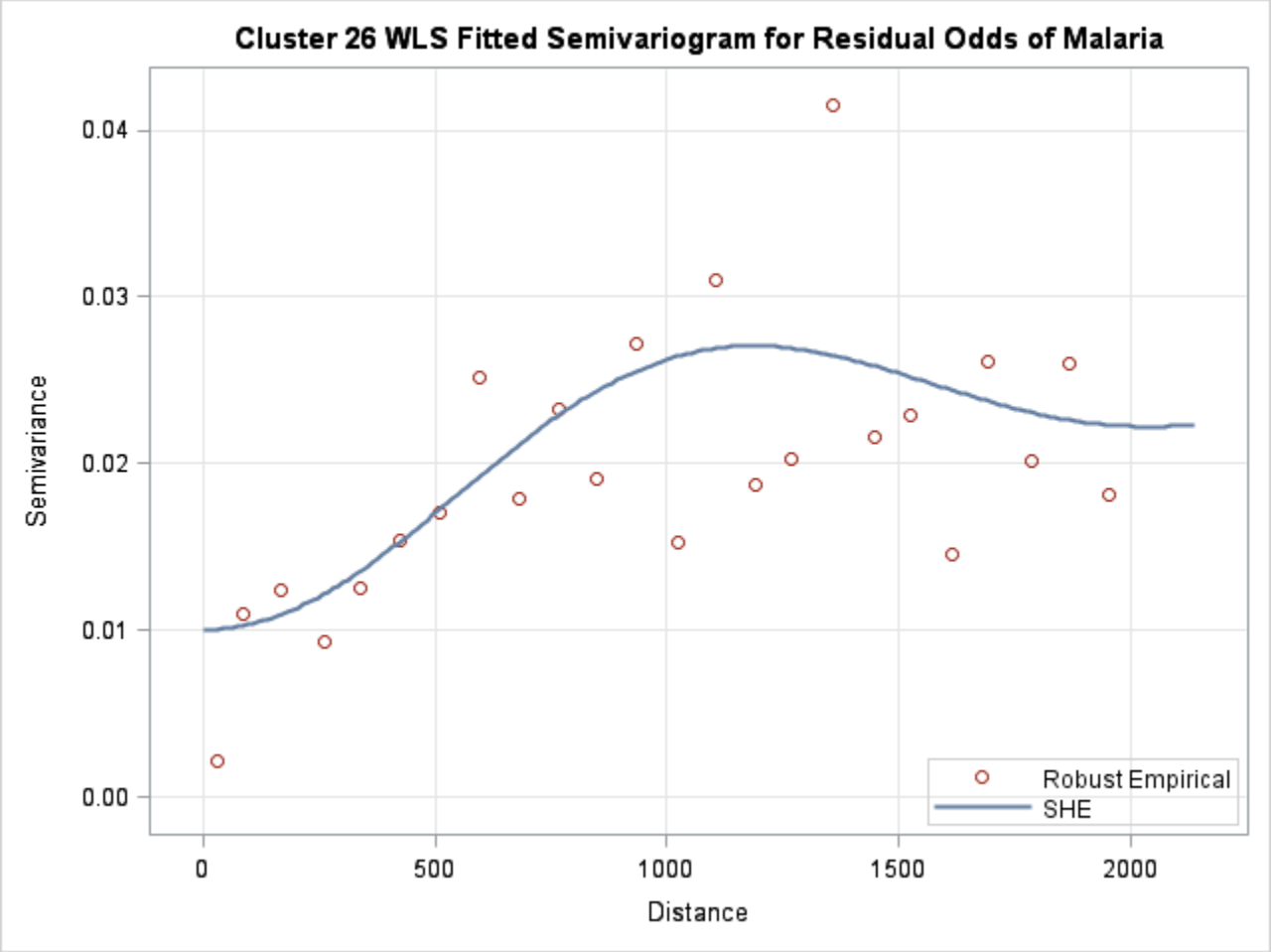


Figure 8.33 *Cluster 27 WLS Fitted Semivariogram for Residual Odds of Malaria*

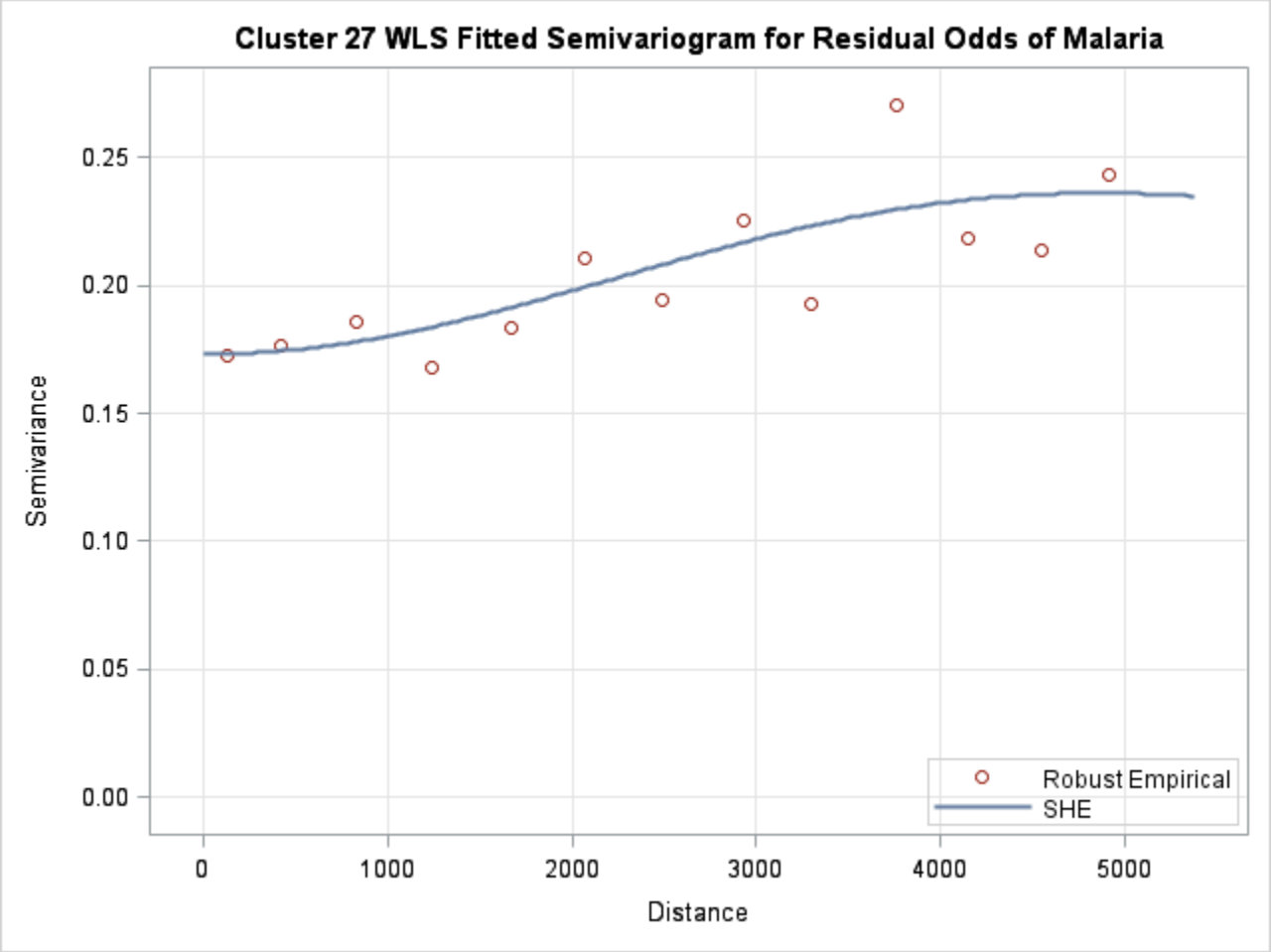


Figure 8.34 *Cluster 28 WLS Fitted Semivariogram for Residual Odds of Malaria*

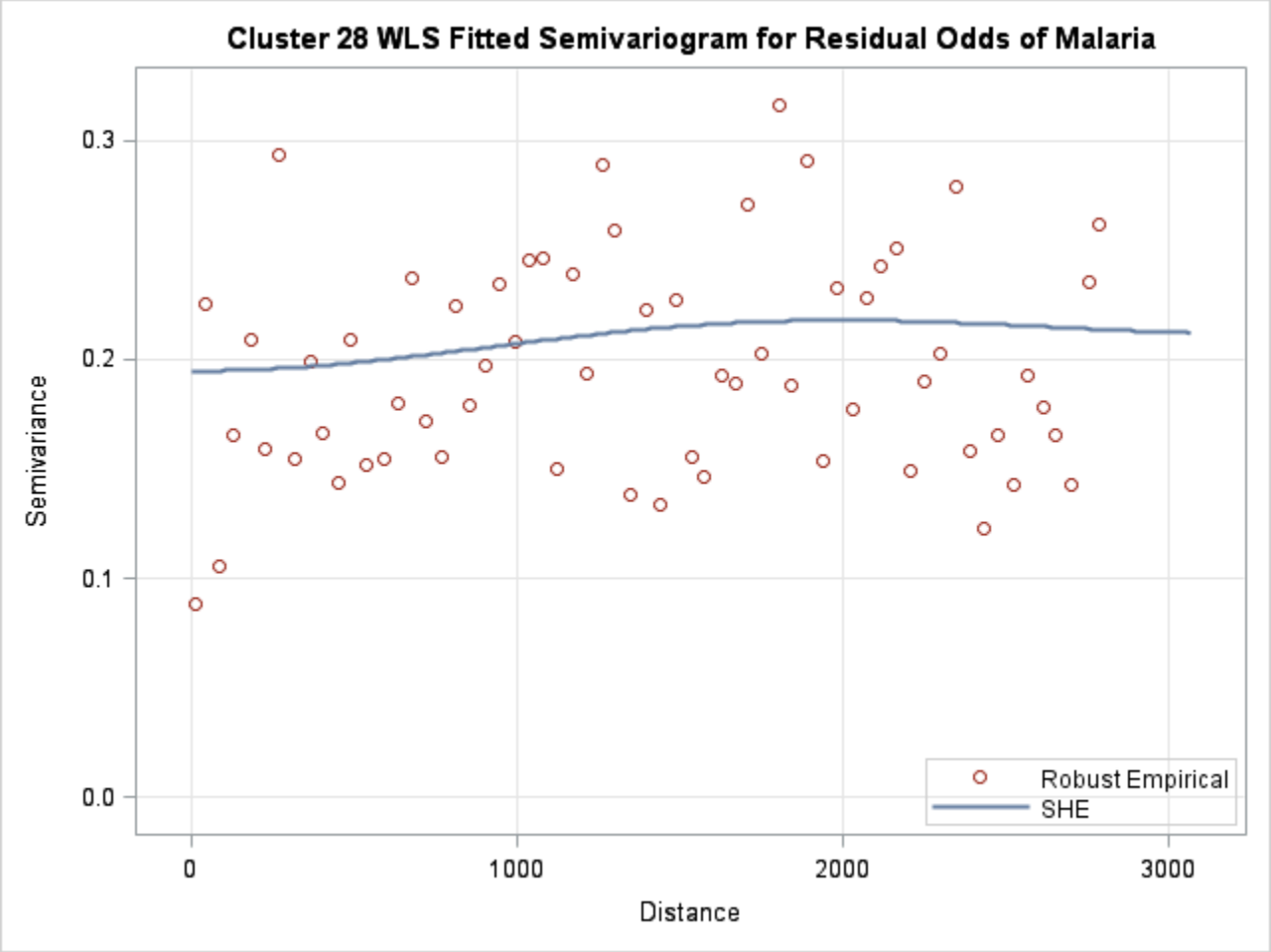


Figure 8.35 *Cluster 29 WLS Fitted Semivariogram for Residual Odds of Malaria*

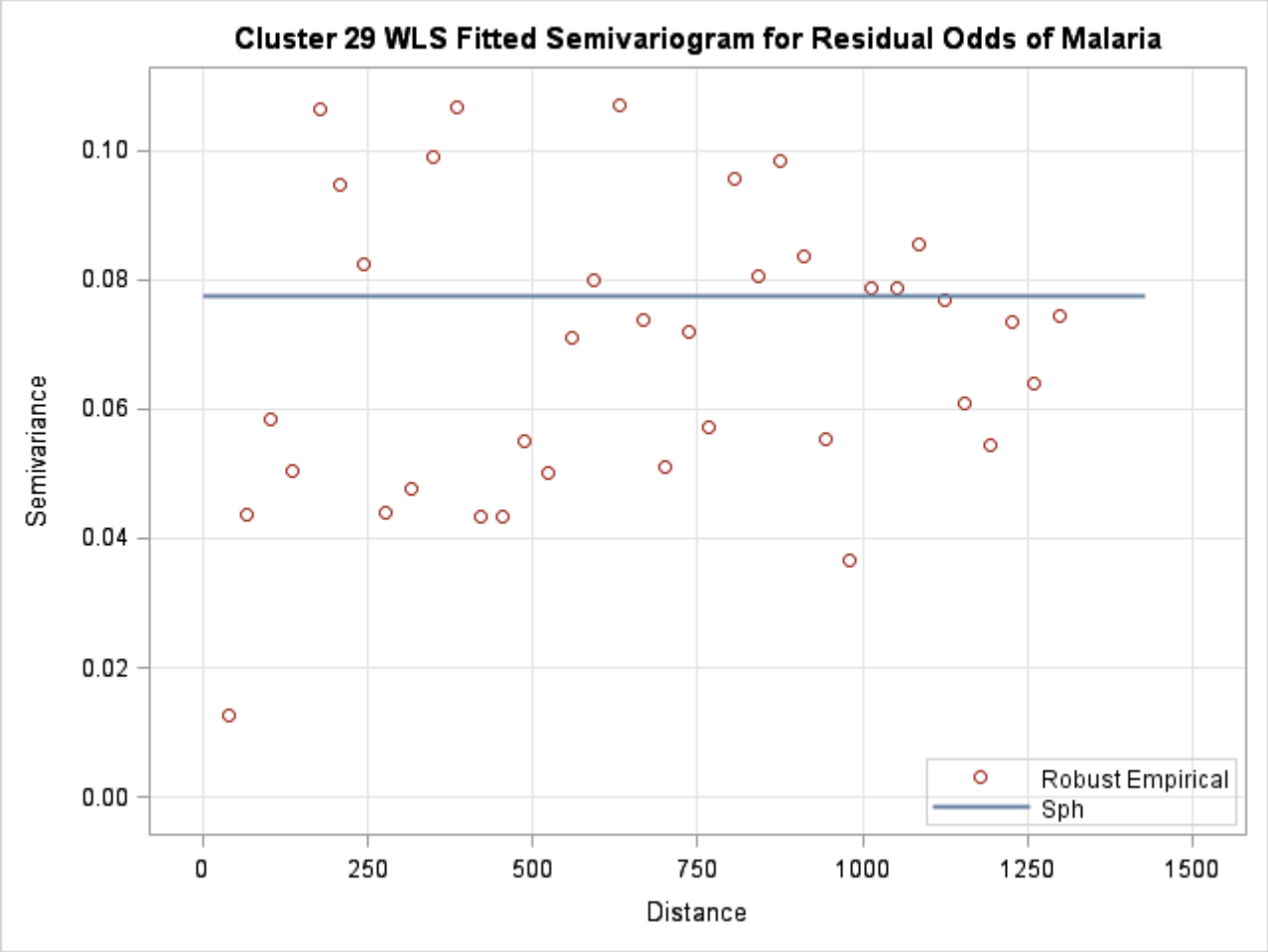


Figure 8.36 *Cluster 30 WLS Fitted Semivariogram for Residual Odds of Malaria*

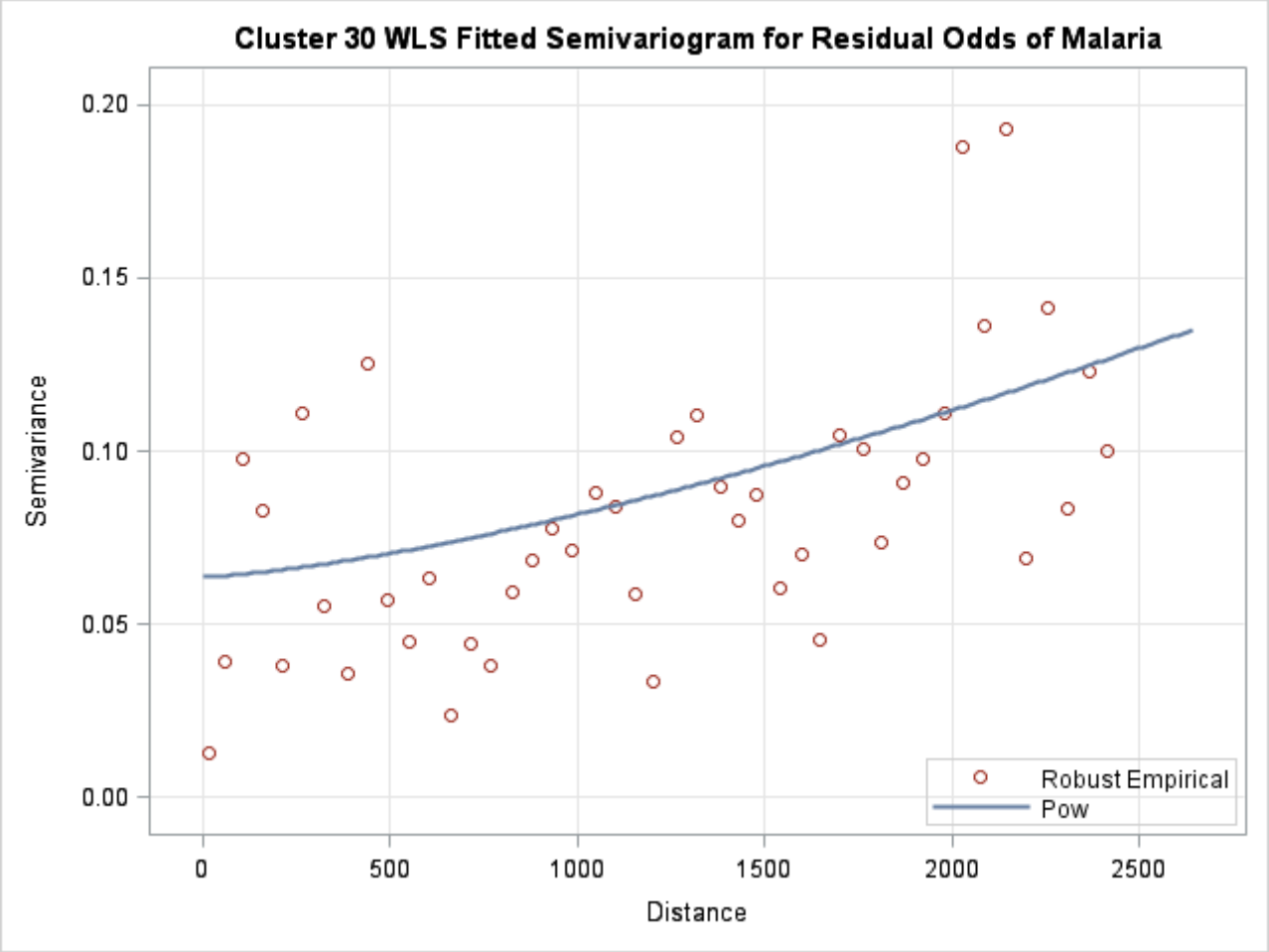


Figure 8.37 *Cluster 31 WLS Fitted Semivariogram for Residual Odds of Malaria*

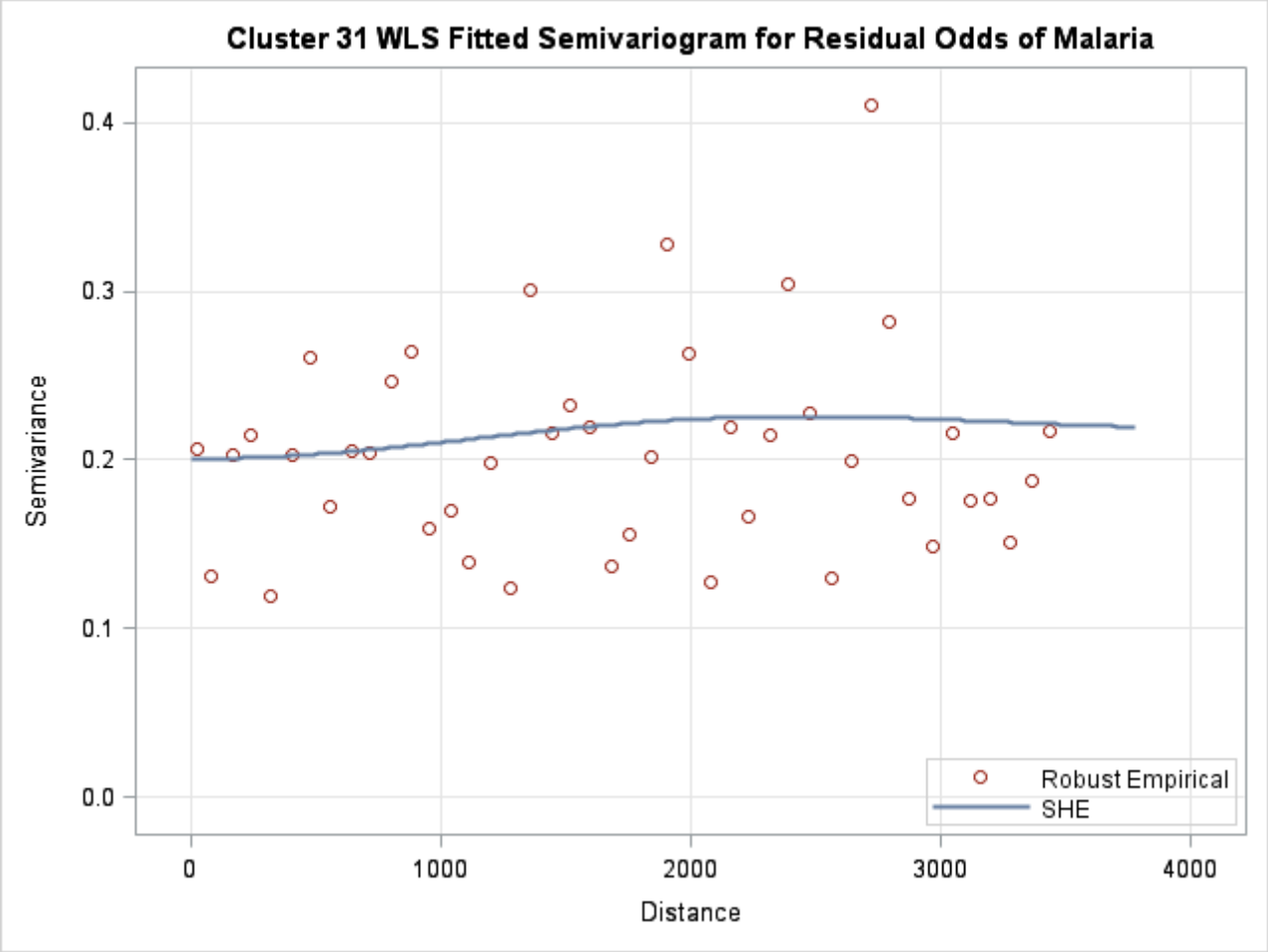


Figure 8.38 *Cluster 32 WLS Fitted Semivariogram for Residual Odds of Malaria*

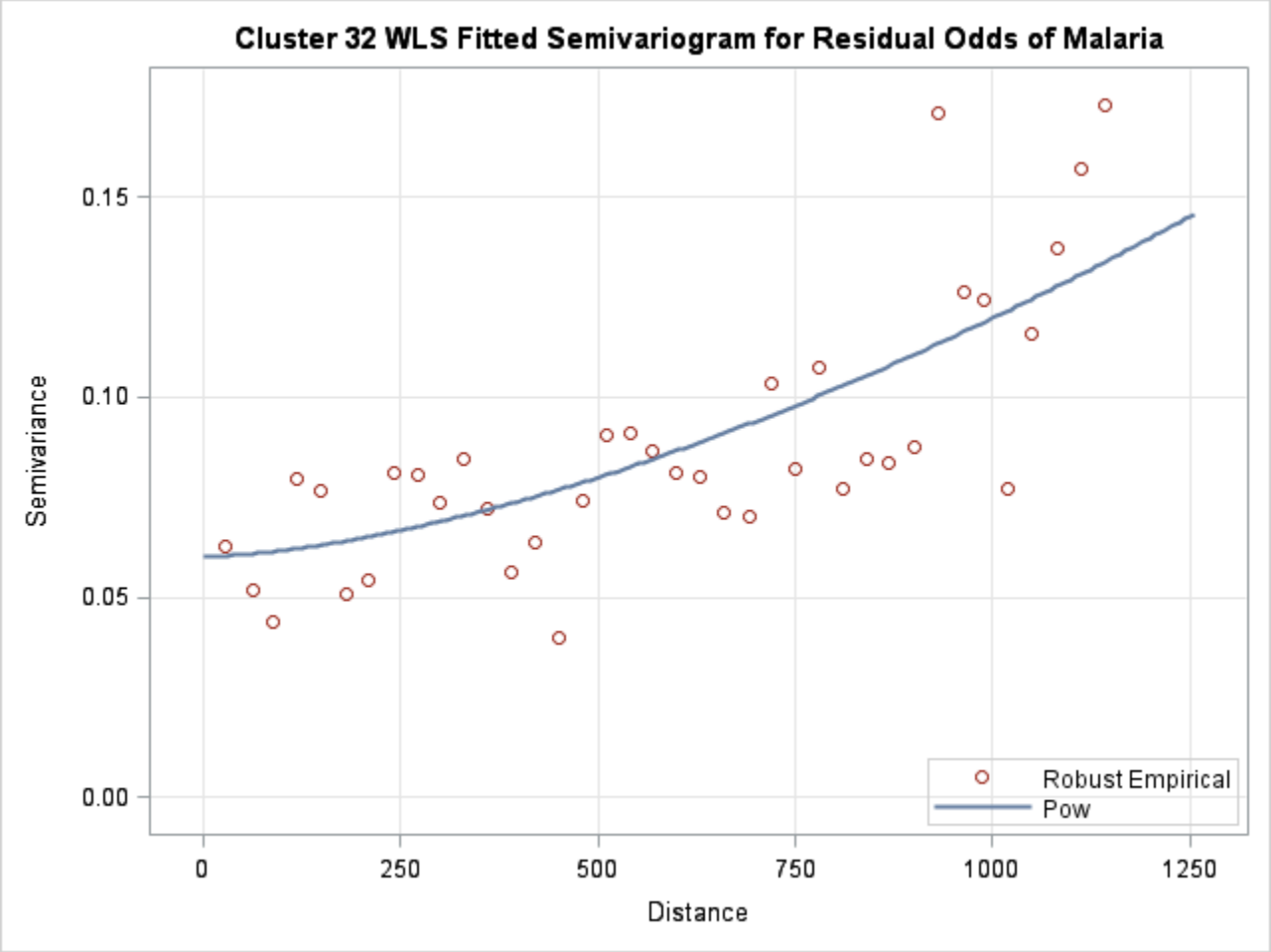


Figure 8.39 *Cluster 33 WLS Fitted Semivariogram for Residual Odds of Malaria*

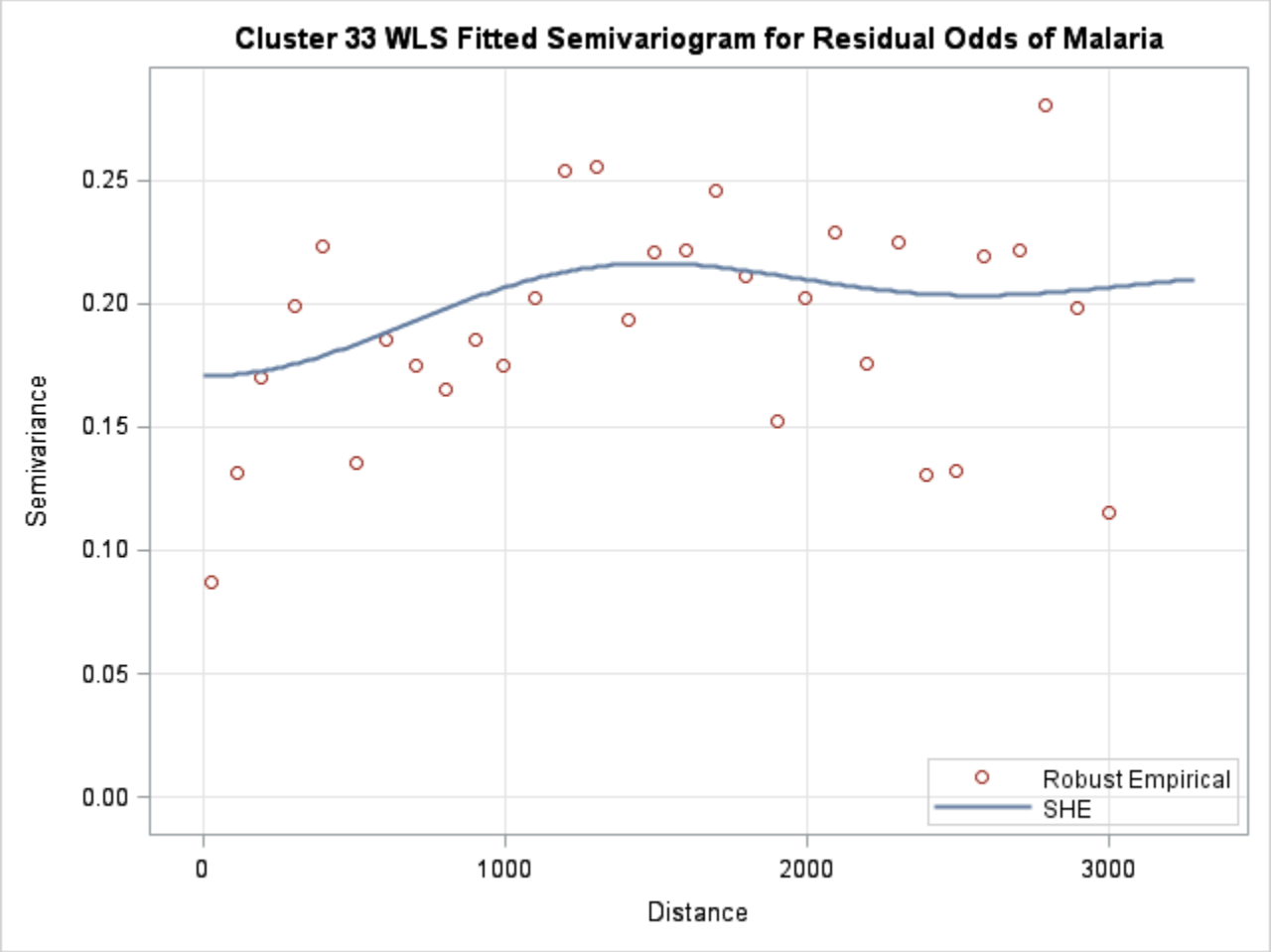


Figure 8.40 *Cluster 34 WLS Fitted Semivariogram for Residual Odds of Malaria*

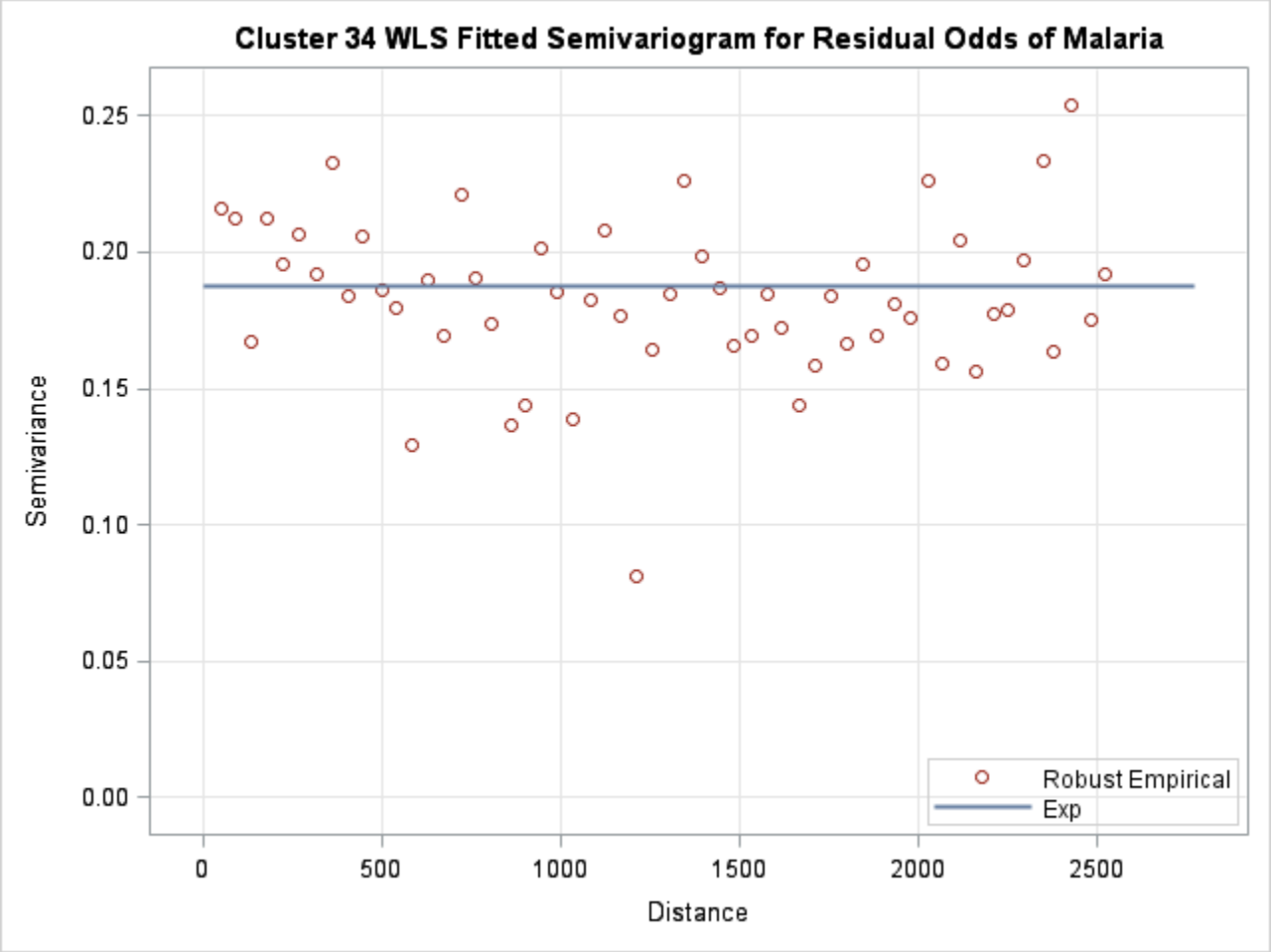


Figure 8.41 *Cluster 35 WLS Fitted Semivariogram for Residual Odds of Malaria*

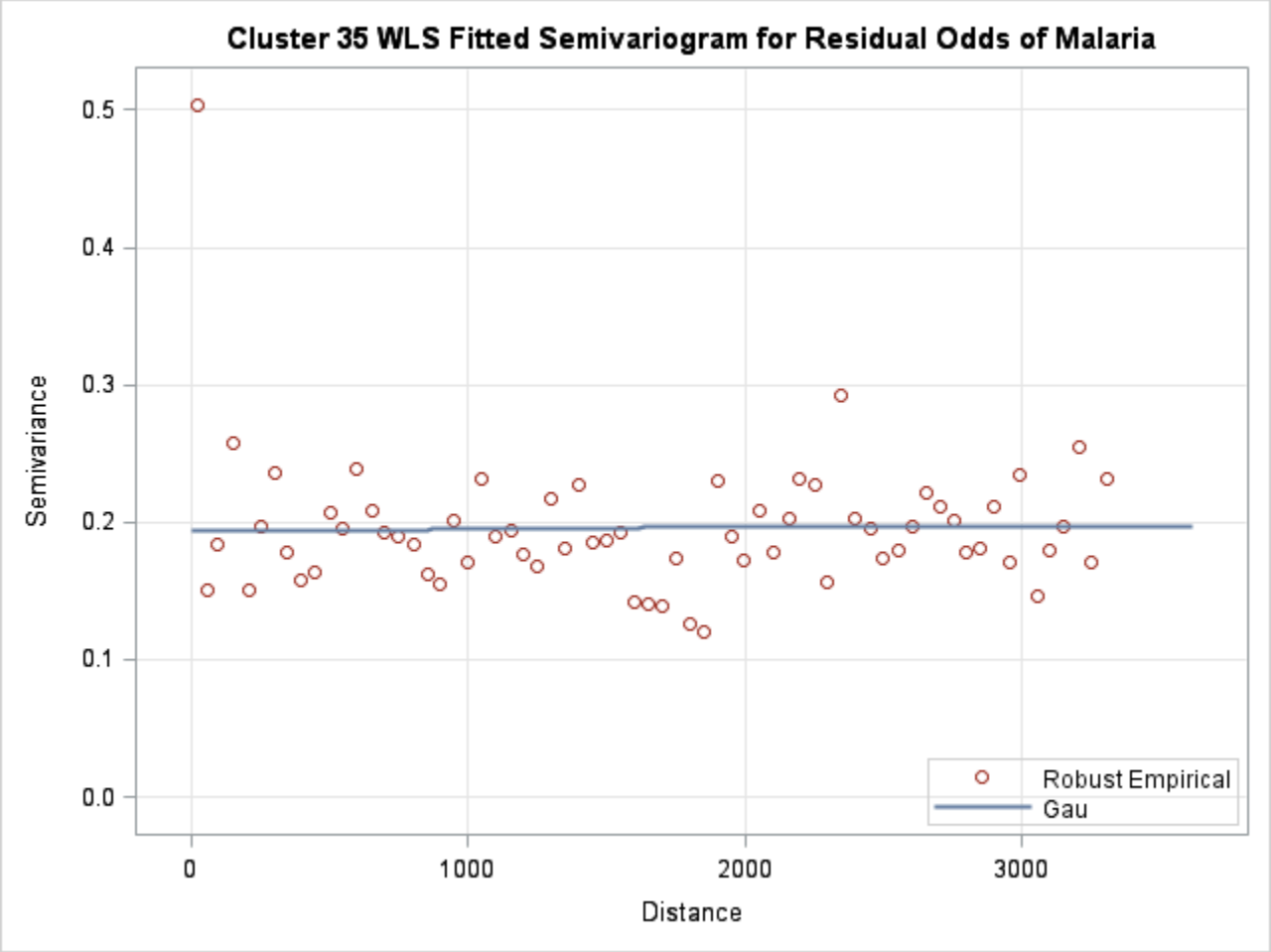


Figure 8.42 *Cluster 36 WLS Fitted Semivariogram for Residual Odds of Malaria*

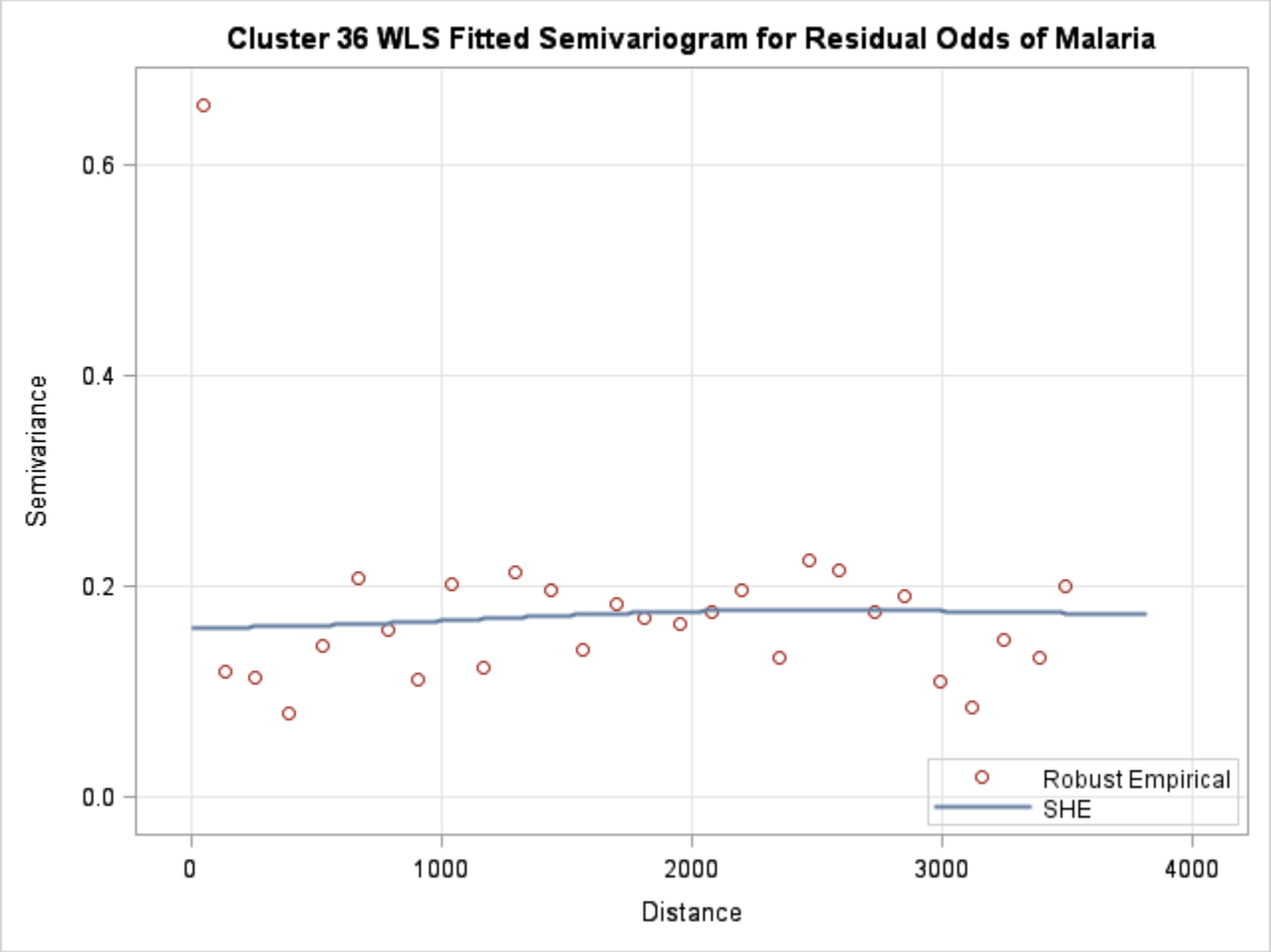


Figure 8.43 *Cluster 37 WLS Fitted Semivariogram for Residual Odds of Malaria*

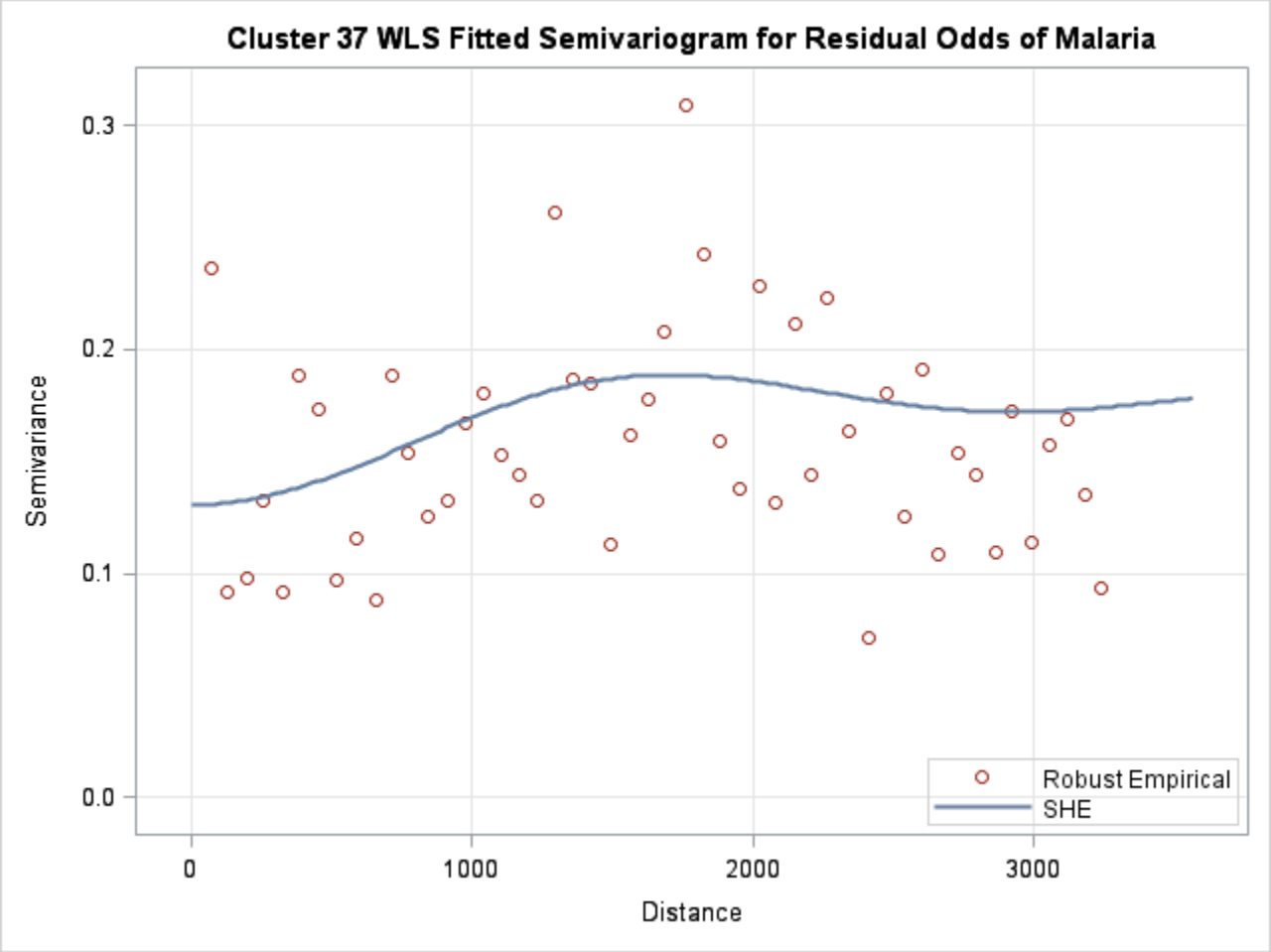


Figure 8.44 *Cluster 38 WLS Fitted Semivariogram for Residual Odds of Malaria*

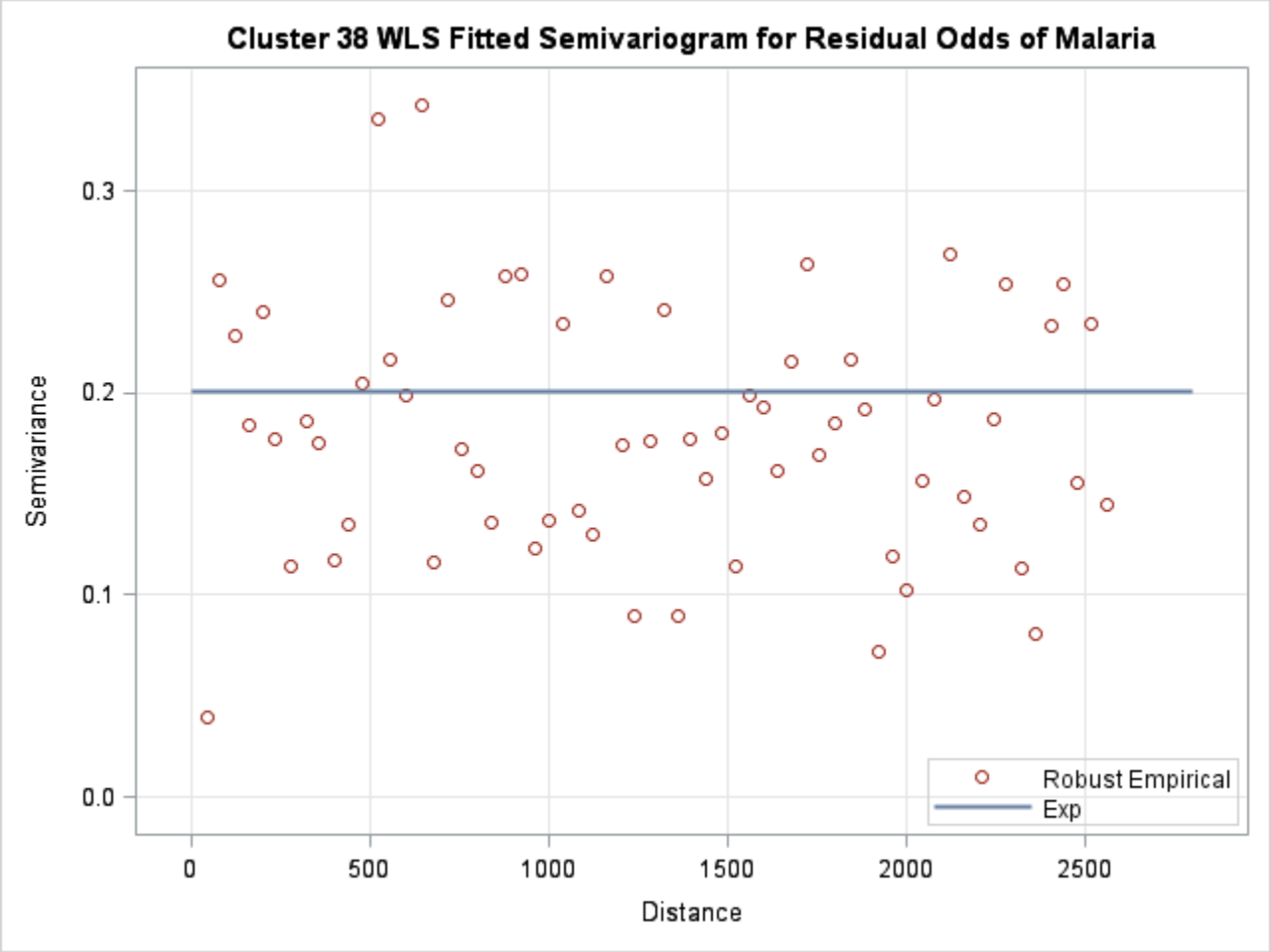


Figure 8.45 *Cluster 39 WLS Fitted Semivariogram for Residual Odds of Malaria*

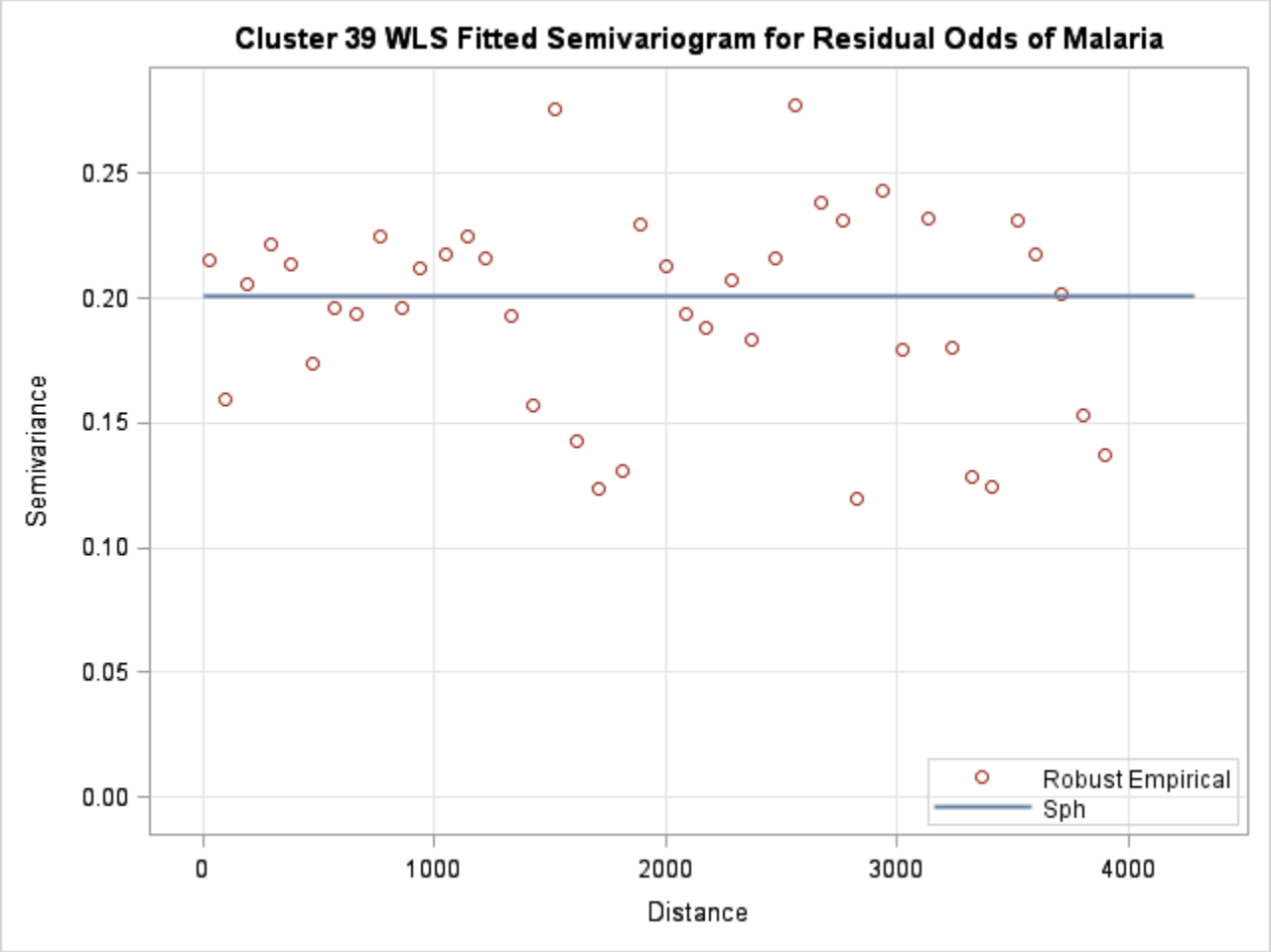


Figure 8.46 *Cluster 40 WLS Fitted Semivariogram for Residual Odds of Malaria*

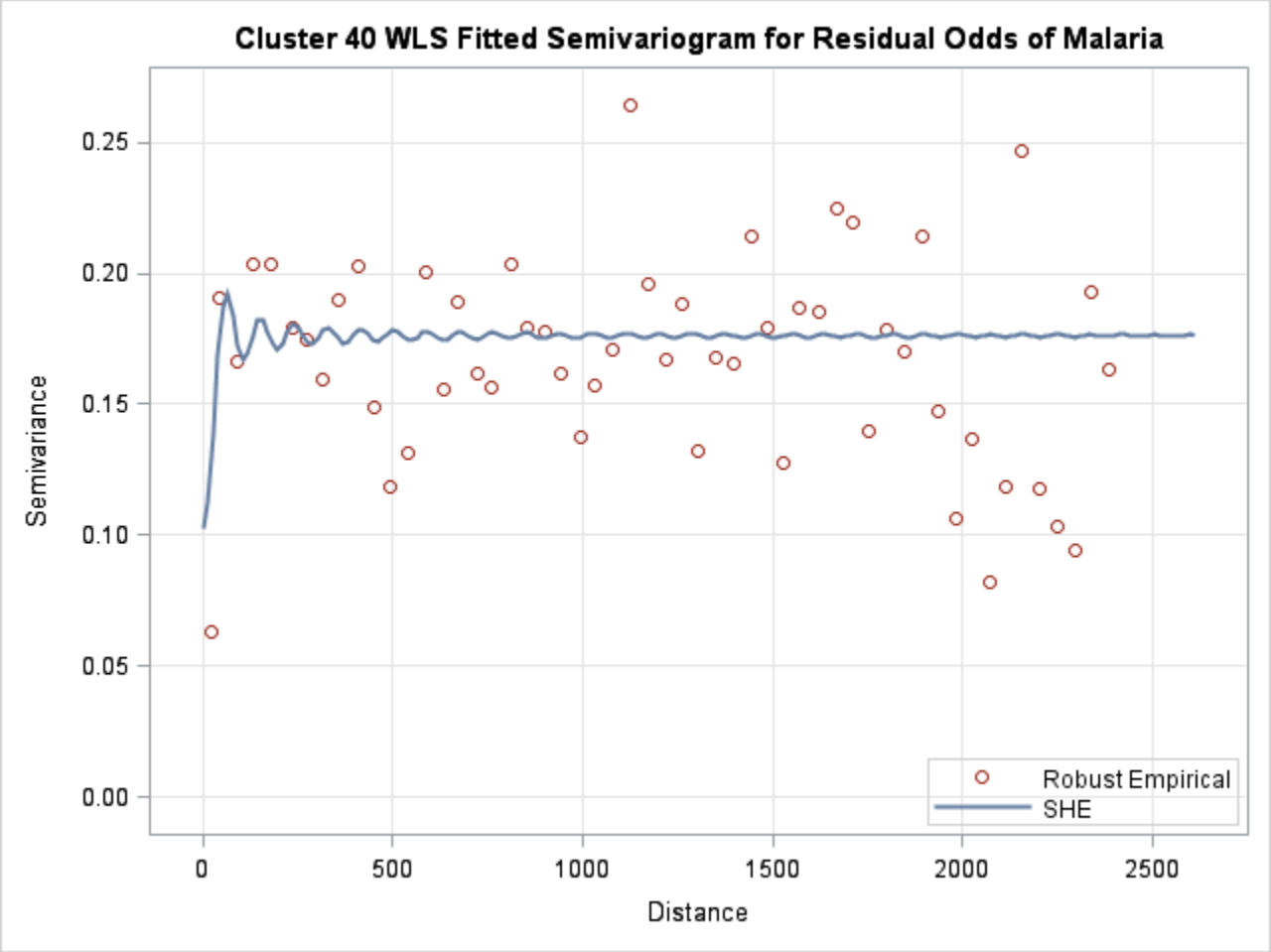


Figure 8.47 *Cluster 41 WLS Fitted Semivariogram for Residual Odds of Malaria*

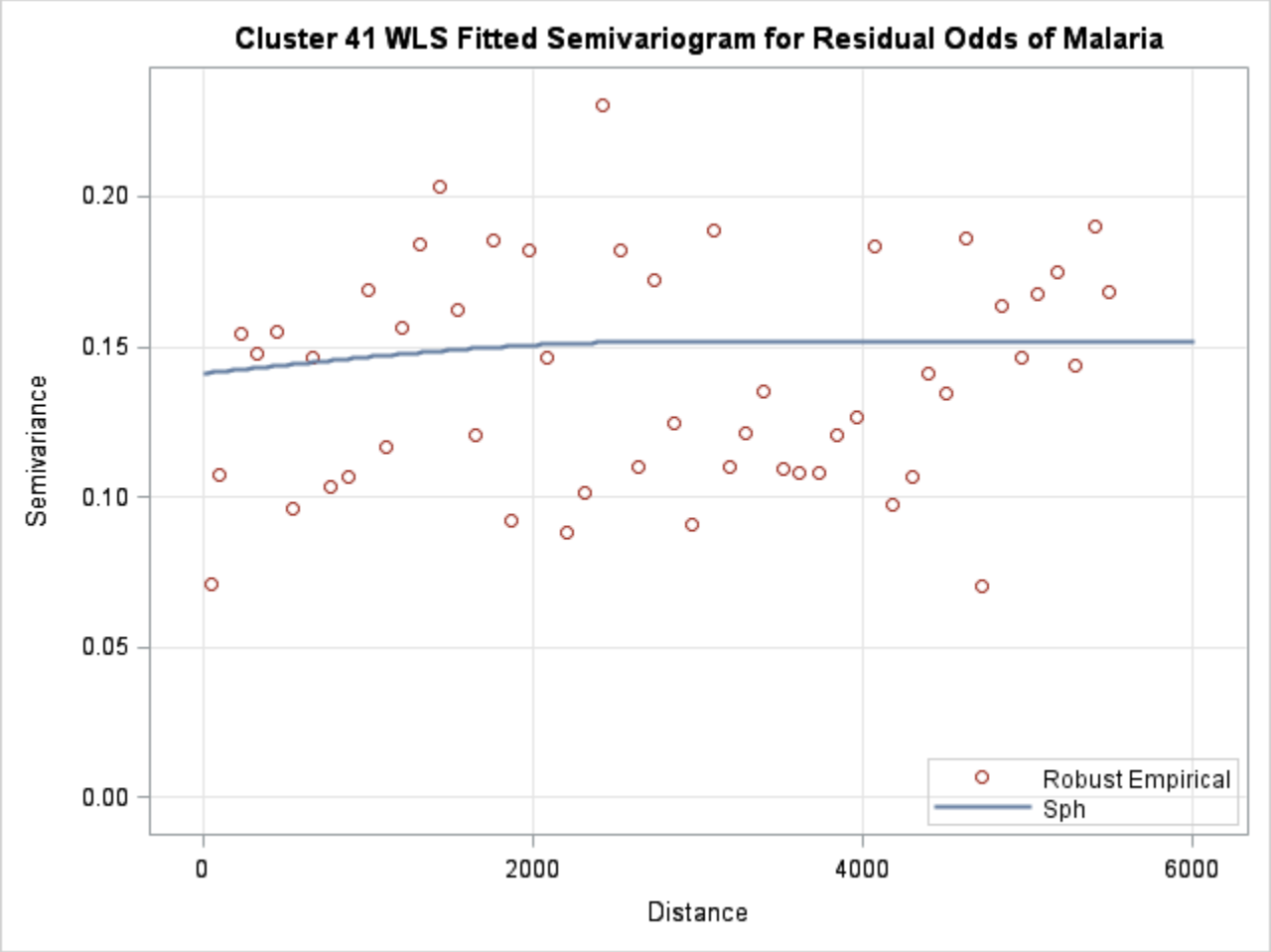


Figure 8.48 *Cluster 42 WLS Fitted Semivariogram for Residual Odds of Malaria*

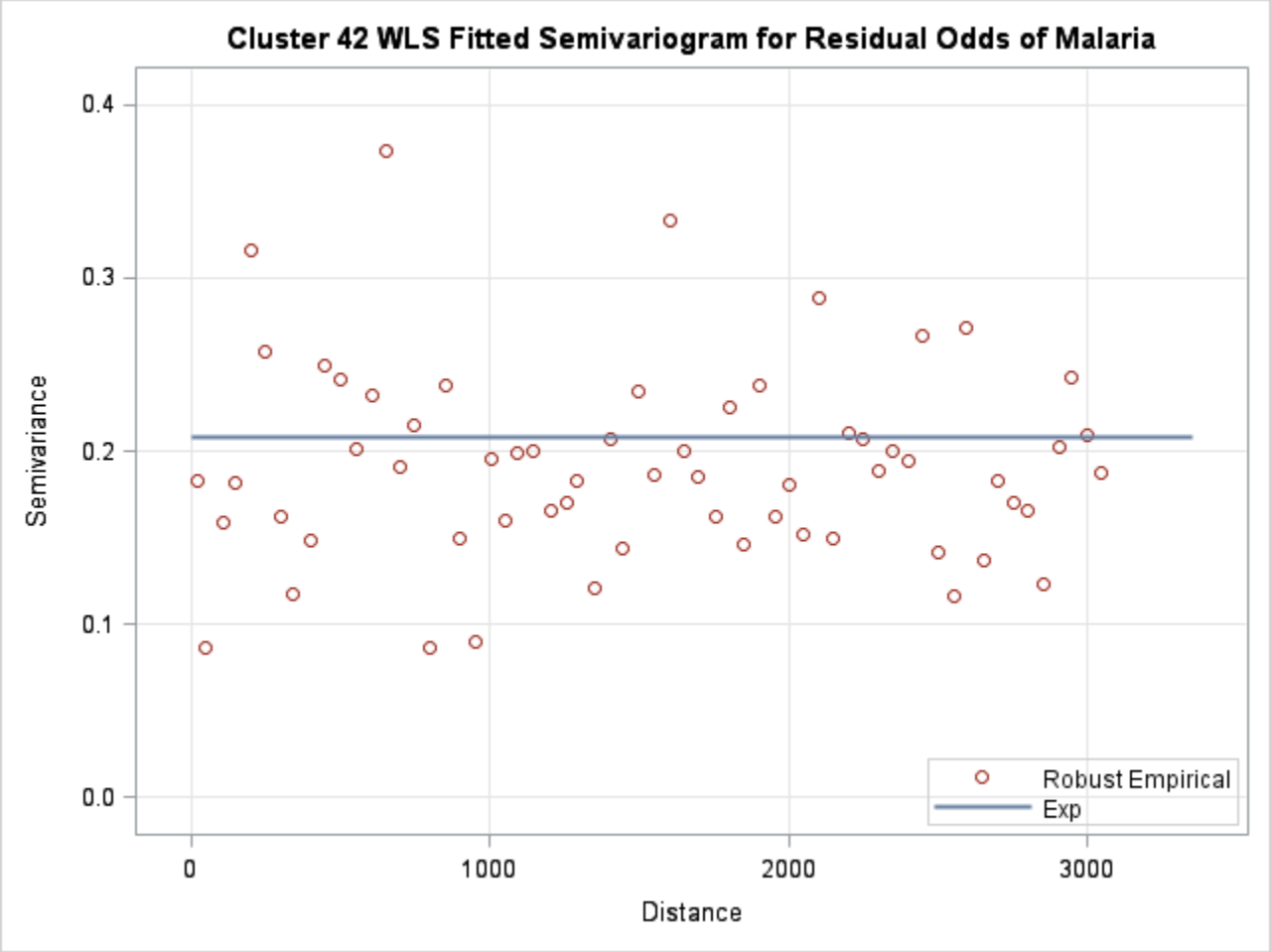


Figure 8.49 *Cluster 43 WLS Fitted Semivariogram for Residual Odds of Malaria*

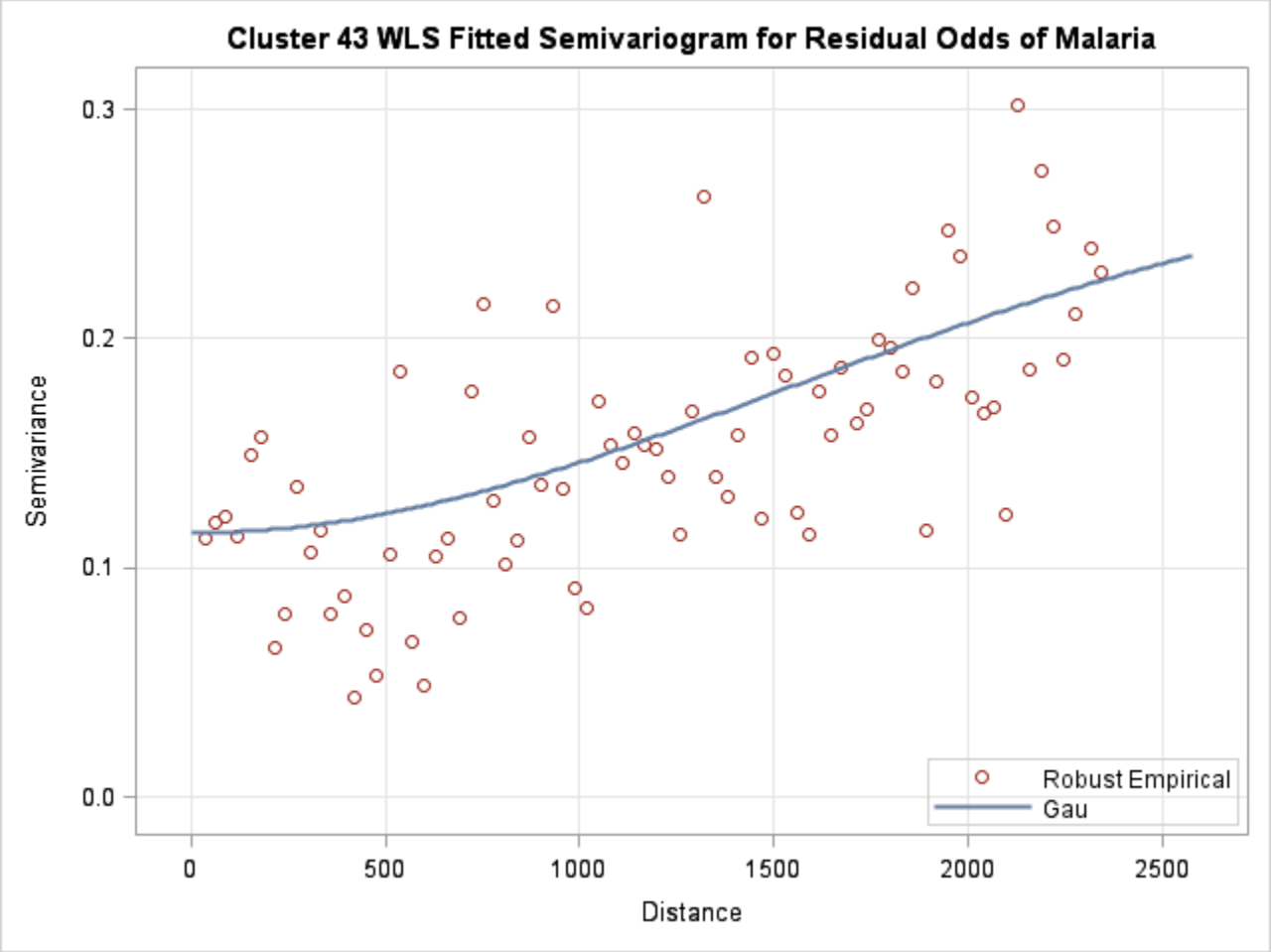


Figure 8.50 *Cluster 44 WLS Fitted Semivariogram for Residual Odds of Malaria*

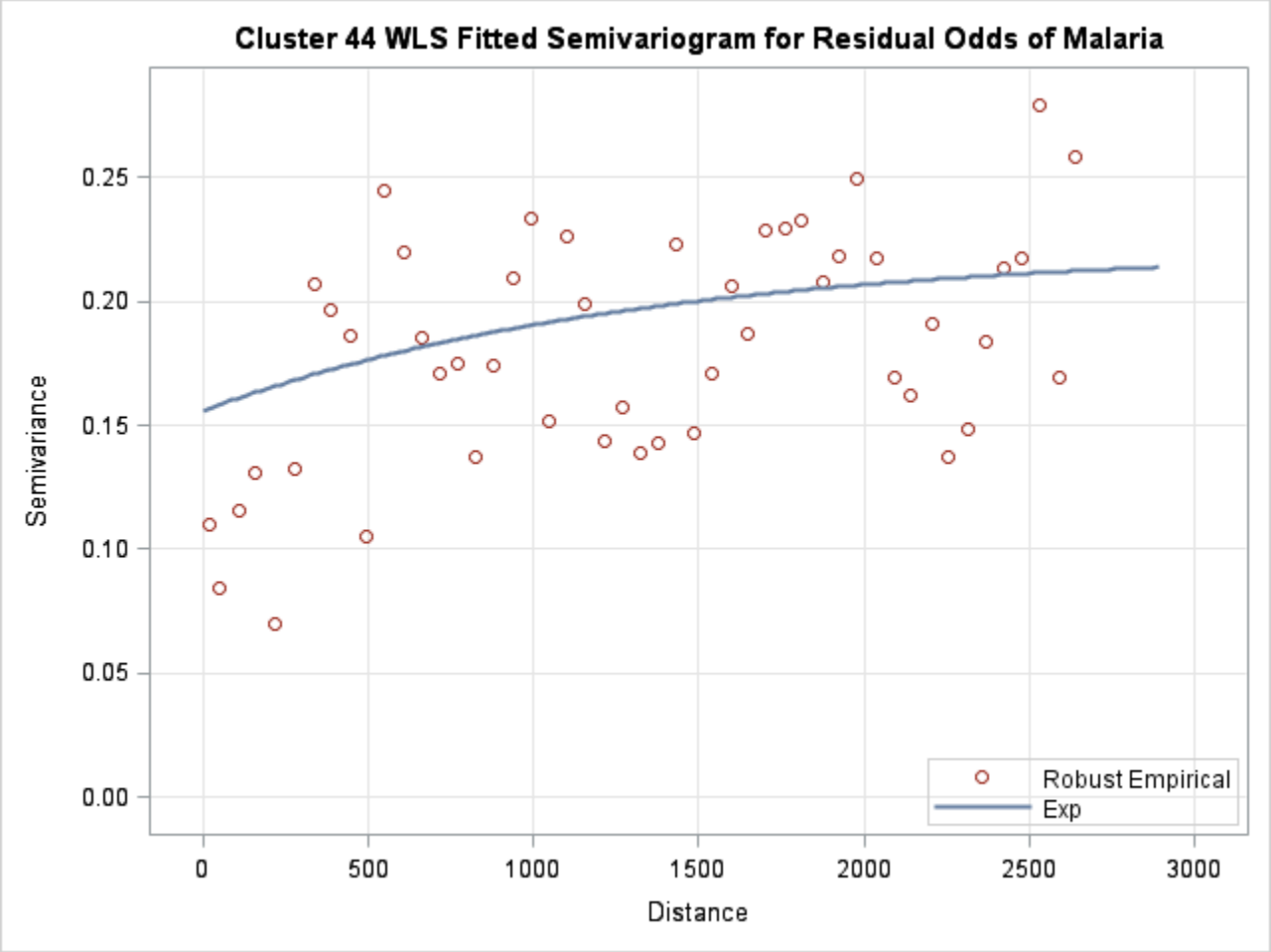


Figure 8.51 *Cluster 45 WLS Fitted Semivariogram for Residual Odds of Malaria*

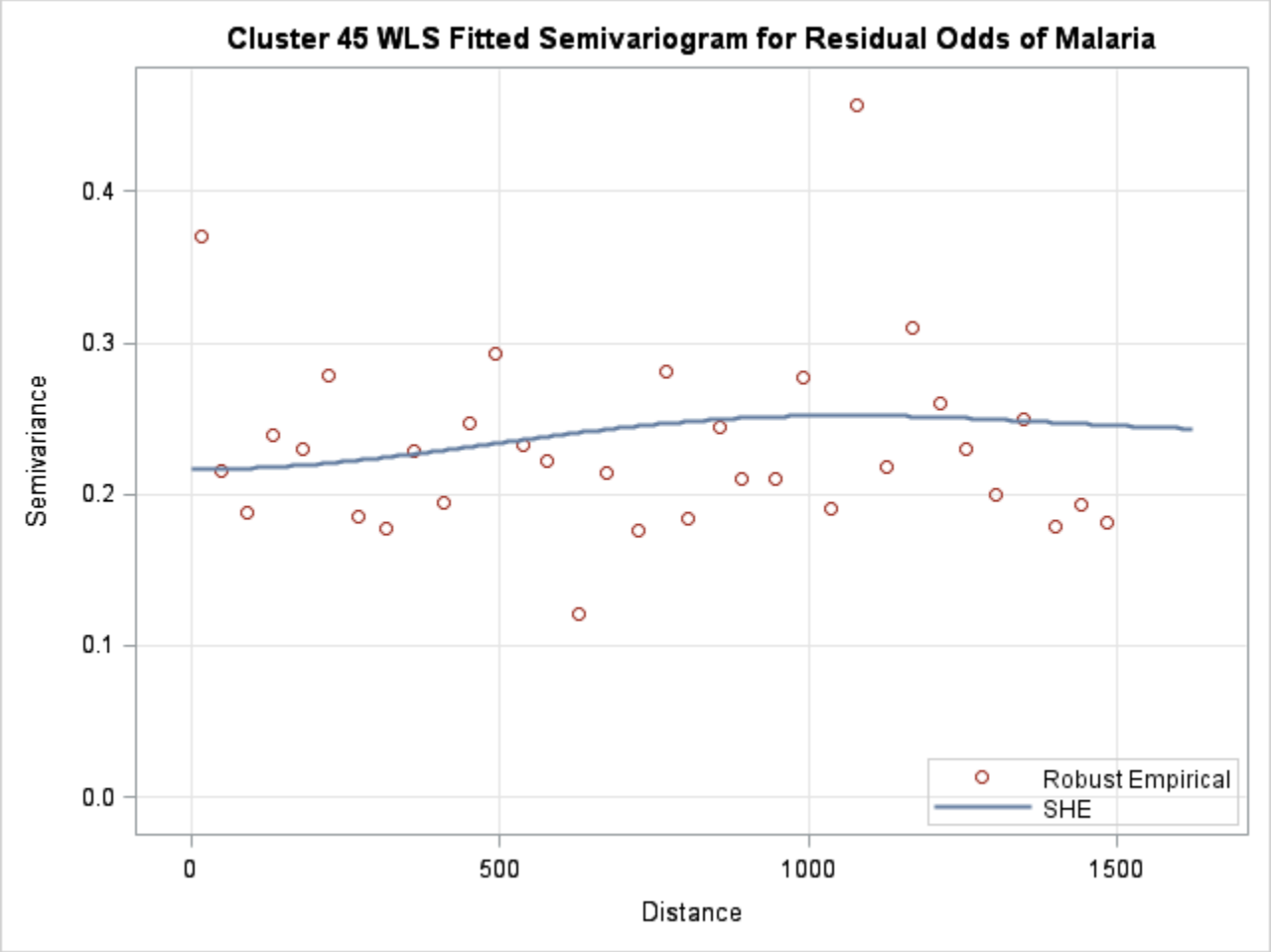


Figure 8.52 *Cluster 46 WLS Fitted Semivariogram for Residual Odds of Malaria*

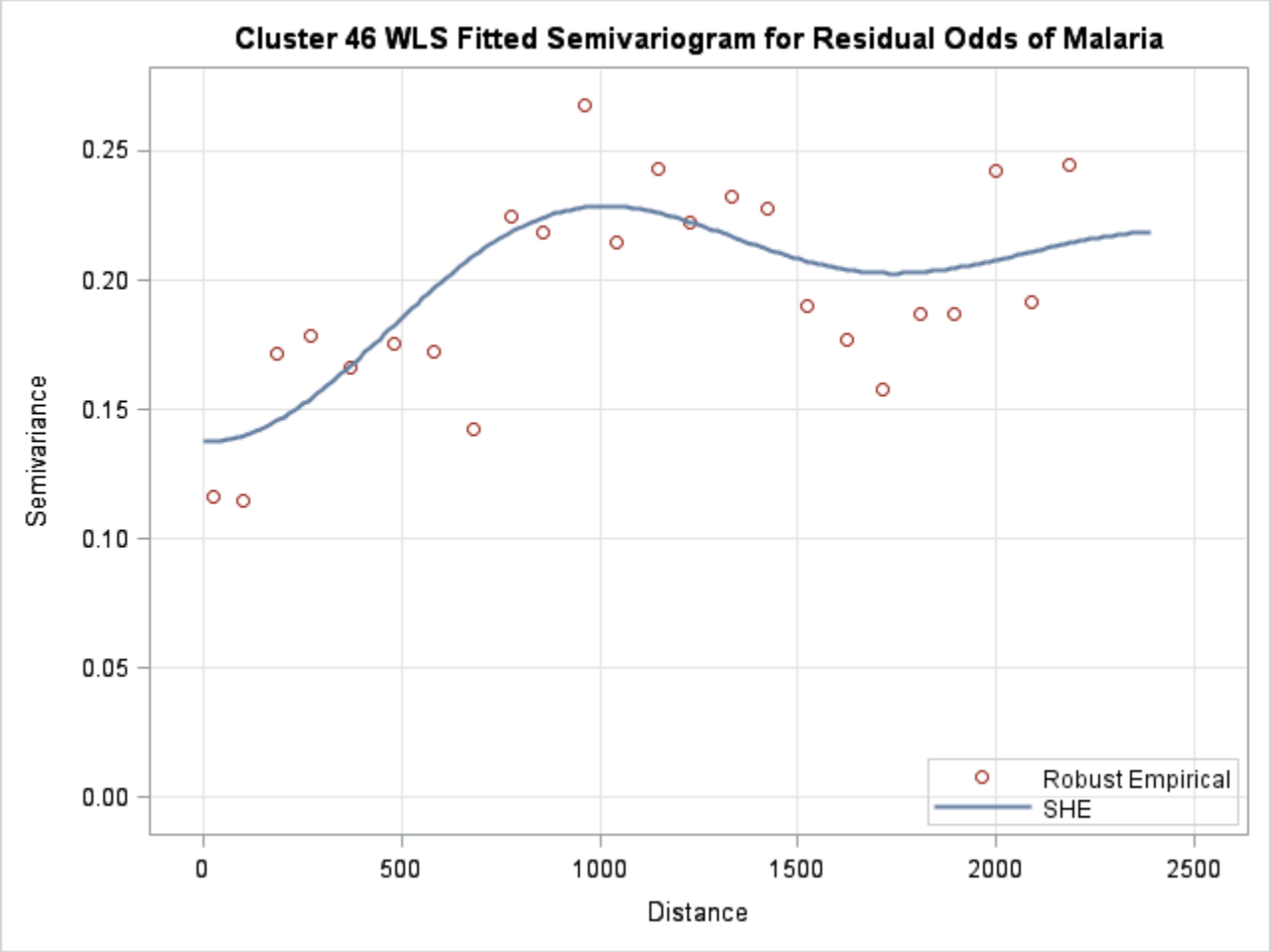


Figure 8.53 *Cluster 47 WLS Fitted Semivariogram for Residual Odds of Malaria*

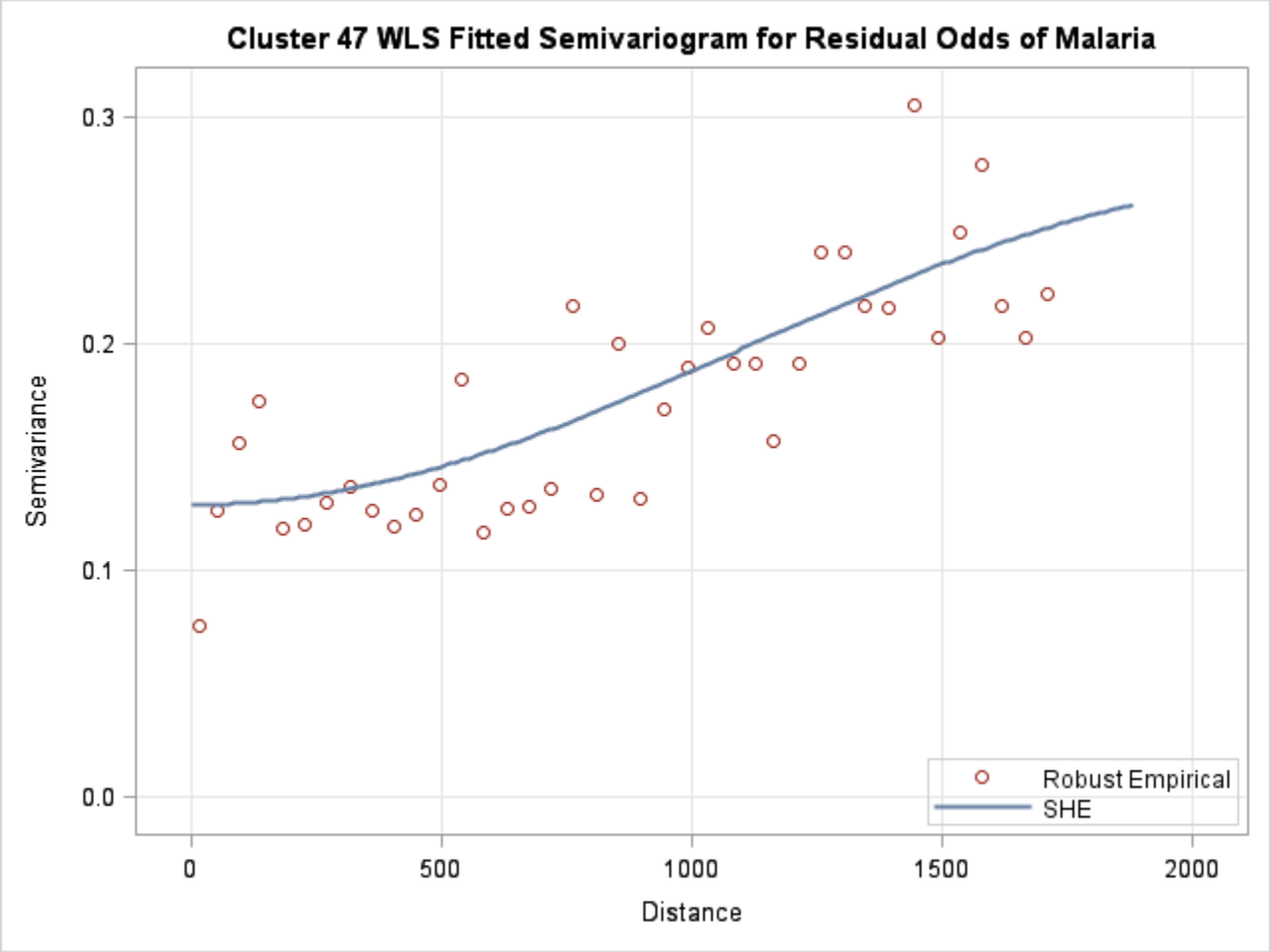


Figure 8.54 *Cluster 48 WLS Fitted Semivariogram for Residual Odds of Malaria*