UNDERSTANDING WETLANDS RECLAMATION AND SOIL-TRANSMITTED HELMINTHS AND SCHISTOSOMIASIS INCIDENCE PATTERNS IN RWANDA (2001-2012)

Elias NYANDWI

Graduation committee:

Chairman/secretary
Dean ITC

Supervisor
Prof.dr.ir. A. Veldkamp University of Twente

Co-supervisor
Dr. S. Amer University of Twente

Members
Prof.dr.ir. A. Stein University of Twente
Prof.dr. M.J. IJzerman University of Twente
Prof.dr. T. Dawson King's College London, UK
Dr. P. Vounatsou Swiss Tropical and Public Health Institute, Switzerland

ITC dissertation number 311
ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN 978-90-365-4410-8
DOI 10.3990/1.9789036544108
https://doi.org/10.3990/1.9789036544108

Cover designed by Benno Masseplink
Printed by ITC Printing Department

UNIVERSITY OF TWENTE.
ITC
FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

UNDERSTANDING WETLAND RECLAMATION AND SOIL-TRANSMITTED HELMINTHS AND SCHISTOSOMIASIS INCIDENCE PATTERNS IN RWANDA (2001 – 2012).

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. T.T.M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 19th October 2017 at 16:45 hrs

by

Elias NYANDWI
born on 5th July 1971
in Ruhango, Rwanda

This thesis has been approved by

Prof.dr.ir. A. Veldkamp, Supervisor
Dr. S. Amer, Co-supervisor

Dedicated to my Family

Acknowledgements

First of all, I would like to express my eternal gratitude to the Almighty God for the extent to which he has enabled me to reach this time alive and to carry out this research.

I would like to express my sincere gratitude to the institutions and individuals, who, in my PhD journey have played an important part towards to the end of this thesis. Please allow me to present my sincere acknowledgments and thanks to the Government of Netherlands through the Netherlands Initiative for Capacity Development in Higher Education under the project NICHE/RWA/071 for financial support provided for this study. To the University of Rwanda (my employer) for facilitating to pursue studies. We thank the Rwandan Ministry of Health through the Rwanda Biomedical Centre for providing epidemiological data and technical support. We are equally thankful to the Rwanda Meteorological Centre, Rwanda Natural Resources Authority, Rwanda Environmental Management Authority and National Institute of Statistics of Rwanda for providing useful data and information for this research.

My deepest appreciations and heartfelt thanks go to Professor Tom Veldkamp, my Promoter, for his excellent advice, guidance, and encouragements. I remember discussing with you, aside from your heavy agenda during your short stay Rwanda, one page, and undecided concept note, and you helped me to sharpen my ideas that eventually led to the present research. From your constructive critiques, your insightful discussion, suggestions and timely feedback, I learned a lot from both my research and career. I would also like to thank my Supervisor Dr. Sherif Amer for your excellent supervision, your enthusiasm, and motivation. You have been supportive during difficult times of my PhD journey. To you both Prof Tom and Dr. Sherif I will ever keep good memories of how you skilfully assisted when my laptop was stolen and lost more than a year of my PhD work.

My gratitude extends to the Faculty of Geo-Information Science and Earth Observation of the University of Twente for creating a multicultural and multi-disciplinary environment excellent as a learning environment. I am very grateful to Dr. Frank Badu Osei. You have been there whenever I needed you, regardless your heavily loaded schedule in a new working position. I will be ever grateful for your contribution in spatial statistical approach and tools for health. Many thanks go to Loes Colenbrander, Petra Weber, Marion Pierik, Saskia and other staff members of the different departments - Library, ITC Hotel, Student Affairs Office, travel unit, help desk and communication for your efforts right from my study registration until the thesis compilation and

layout. I can’t thank you enough for your onsite and remote support in handling various issues which made my stay at ITC pleasant.

I am very grateful to all those who made valuable contributions during my stay in Netherlands, fieldwork in Rwanda and somewhere else during scientific conferences and seminars participation: Dr Irenée Umulisa (RBC/M&OPD Department); the International Society of Geospatial Health which I knew and adhered to during my PhD journey; Prof Arnaud Temme (of Wageningen University – Soil and Landscape Group). Thanks to fellow PhD Students Alphonse, Adrie, Marie Christine, Clarisse, Prosper, Vestine, Alice, John, Matthew, Ibrahim, Festus, Omar, Kapongola, Katomero, Tsega, Ullah, Yusof, Vincent, Zaid. To my officemates Divyani and Rashidi, I owe you special thanks. You have been supportive and encouraging. I would like also to acknowledge support from the family of Vincent Briac & Anne Sonet, Olivier Nzamuye, Claudine and others families and friends at Enschede, Netherlands, and Belgium. My profound appreciations to the Rwandese Master's students between 2012- 2016 with whom we shared time in Enschede. Many thanks to Frodouald, Haguma, Alphonse K., Theodomir, Mireille, Felicien, Herve, Joseph, Sylvain, Potel, Guillaume, Francois, Emmanuel I., Gilbert, Appolonie, Ignace, Marc, Fred, Dominique, Maurice, Jean Maurice, Eric, Honoré, Providence, Jean Baptiste, Marceline, Aline, Ahmed, Crispin, Janvier, Innocent, Adeline, Aphrodis, Oscar, Emmanuel N., Celestin and Chantal.

My academic destiny began on that first day in primary school. I dedicated this work to my late parents, to my brothers and sisters, extended family members and friends for your support. Without listing all the names please accept my sincere thanks for your prayers. My special thanks and gratitude are due to my lovely wife Ida Nibamuhoze, and my sons Ndinda Vertex Noûr and Ndaje Brave Naïmu. I am highly indebted to you for approving doing my PhD by foregoing the right to be cared by a husband and father, *muri Intashyikirwa*. I am delighted to dedicate this work to you.

Table of contents

List of Figures

List of Tables

Chapter 1: General Introduction

1.1 Background

Within-wetland ecosystems people try to improve their livelihood by converting land cover, extracting resources and redirecting water flows. From these human actions, several intended or unintended positive or negative environmental effects are induced. The benefits or positive impacts of wetlands development projects include, amongst others, food production and fresh water supply. While intending to maximize ecosystem benefits, the modified environment can also have negative impacts. One of these negative effects in Sub-Saharan Africa is an increased of risk of transmission of schistosomiasis and other water/soil-transmitted helminths to humans. Therefore, understanding the relationship between water/wetland environment change and its potential impact on diseases distribution is essential for sustainable human development. In this chapter, background information is provided on the concept of wetland environment characterisation, wetland values and their link with soil-transmitted helminths (STHs) and schistosomiasis. It is also explained why and how the link between wetland ecosystems and STHs and schistosomiasis need to be better understood. Finally, the objectives and structure of this thesis are presented.

Wetlands definition, characterization and value

Ecosystems are complex and rapidly impacted by human activities. With increasing population and impacts, not only the health of the ecosystems that support us can be in peril, but also human health itself (2012). Wetland ecosystems are lands where water is the dominant factor determining the nature of soil development and the types of plant and animal communities living in the soil and on its surface (Cowardin, Carter, Golet, & LaRose, 1979). As illustrated in Figure 1, wetland characterization involves three main components: presence of water (either at the surface or within the root zone), unique soil conditions that differ from adjacent uplands, and vegetation adapted to wet conditions (hydrophytes) with absence of flood-intolerant vegetation (Hollis, 1993; Mitsch & Gosselink, 2000). To establish a sustainable use of wetlands, various national and international efforts supported by the Ramsar Convention for accurate mapping, inventory, monitoring and assessment of wetlands characteristics have been developed (Ramsar, 2007).

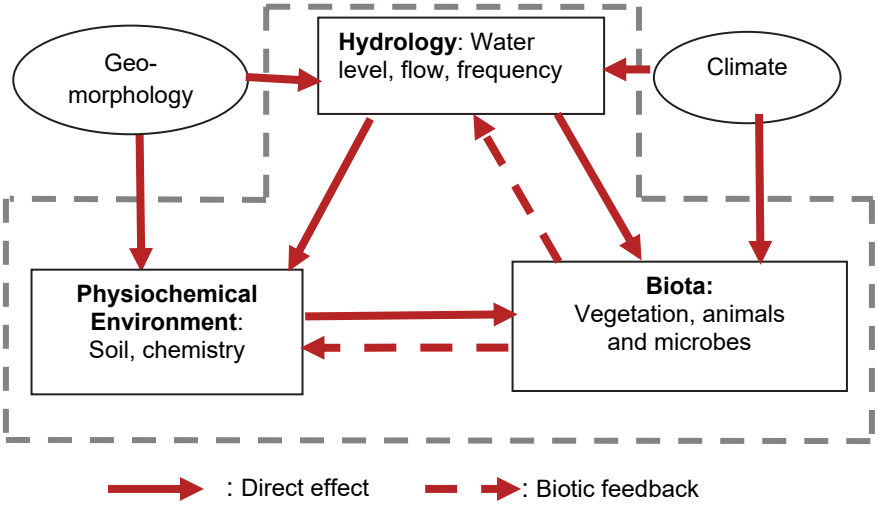

Figure 1: Elements of a wetland. The diagram summarizes the value of wetlands concept, as adapted from Mitsch and Gosselink (2000).

These three elements are not independent. Climate and geomorphology define the degree to which wetlands can exist, but the starting point is the hydrology, which affects the physicochemical environment (soils), which in turn determines what and how much biota (flora and fauna) is found in the wetland.

In this context, the European Space Agency (ESA) in collaboration with the Ramsar Secretariat launched in 2003 the "Global Wetland Geo-information system" project using Earth Observation technology to support inventorying, monitoring and assessment of wetland ecosystems involving 50 different wetlands across 21 countries on four continents (Jones et al., 2009). Landsat TM images and digital classification techniques were used for delineating western coastal and saline wetlands in Sri Lanka (Rebelo, Finlayson, & Nagabhatla, 2009); and for mapping at a broader scale inland wetlands in Southern Africa using Decision-tree Classifiers on Landsat TM and Aster images (Ewart-Smith, Ollis, Day, & Malan, 2006). The satellite imagery based techniques were good in land use and land cover mapping, but involves significant field visits and/or manual interpretation of aerial photographs, or other very high-resolution earth observation data, to increase their accuracy (Feng & Bajcsy, 2005; Hengl, Gruber, & Shrestha, 2003). Furthermore, Digital Elevation Models (DEM) and their derivatives (slope, curvature, topographic index, etc.), generated easily with standard GIS software, are extremely useful for wetland boundary delineation. Several studies successfully used topographic variables derived from a DEM for automatic delineation of wetland boundaries. In Canada a DEM was used to discriminate

between wetland and upland (Hogg & Todd, 2007). In South Africa DEM's were used to distinguish between real depression features and spurious ones (Temme, Schoorl, & Veldkamp, 2006). DEM's were also successfully applied in modelling the Amazonian floodplain (Yamazaki et al., 2012), and for wetland land use changes due to agriculture and related erosion and deposition (Kassawmar, Rao, & Abraha, 2011; Neumann et al., 2011). However, climate-related data that determine water availability and the local net water balance (such as rainfall and evapotranspiration) are less used for wetland modeling (Adam, Mutanga, & Rugege, 2010). As illustrated in Figure 1, improved modeling of wetlands should consider the interactions between topographic, hydrologic and climate variables.

The environmental services provided by wetlands can be divided into products and functions. Products are tangible outputs of a wetland and can be used directly or sold: agricultural products, fish, wood, honey, handcraft material, etc. Functions are what wetlands contribute to the wider environment and are more indirect in nature: e.g. water filtering, water supply, flood control, groundwater recharge, habitat for wild animals and birds, tourism and recreation (Heady, 1992).

Worldwide, wetlands sustain large communities of people, who depend on their natural resources that maintain them. Indigenous systems of water resource management include agriculture (food cropping, notably of rice, flood recession cropping), fishing and pastoralism (Adamas, 1993). Since the 1970's several agricultural development programs were initiated in inland valley swamps, larger inland deltas and lacustrine wetlands in Western Africa (Adamas, 1993) and Eastern Africa (Jalloh, Roy-Macauley, & Sereme, 2012). Nowadays, intensive and dynamic wetland cropping mainly concerns rice cropping systems. Production practices range from very primitive to high input systems (Beighley, 2012), in naturally flooded or irrigated fields.

In Rwanda, the use of wetlands for intensified agricultural production is a relatively recent phenomenon as indicated by the irrigation master plan of Rwanda (Malesu et al., 2010). The need for increasing agricultural production in Rwanda to achieve MDG1 (eradicating extreme poverty and hunger) is further reflected in wetland development as one of the major activities for the transformation of agriculture from subsistence to market-oriented farming, sometimes despite the regulatory framework defined by the Rwanda Environmental Management Authority (REMA) for sustainable wetlands management. Wetland reclamation increased from approximately 400 ha for tea and rice cultivation in the 1960's to more than 12,000 ha of swampland reclaimed for agriculture in 2011. In the framework of Vision 2020, a long-term strategy for national development, and the Economic Development and Poverty Reduction Strategy (EDPRS), rural development and agricultural

transformation are spearheads for rapid and sustainable development. The growth rate assigned to the agriculture sector is between 5 to 8% each year to reach expected objectives by 2020 (Ministry of Agriculture and Animal Resources, 2004). Wetland conversion is one of the major foreseen activities of intensification and development of sustainable agricultural production systems. Following this policy, at least 40,000 ha of wetlands should be reclaimed in 2020 (Malesu et al., 2010). Indeed, as illustrated in Figure 2, most of Rwanda's wetlands are being reclaimed under government schemes with the aim of growing rice as main crop (Nabahungu & Visser, 2011). Three different cover types can be observed in Rwandan wetlands: (i) wetlands with natural vegetation, under protection or inaccessible (with deeper water during a larger period of the year); (ii) wetlands with a mixture of natural vegetation and crops, and (iii) wetlands completely and/or permanently used for agricultural activity (Lozano et al.).

Figure 2: Rice farmers in Rongi sector of Muhanga district. Around a newly reclaimed wetland for rice cultivation - the involved community gathers around the agriculture officer and president of farmers' cooperative for guidance.

Wetland conversion, schistosomiasis and soil-transmitted helminths

Intestinal Helminths are parasitic infections which cause thousands of avoidable deaths each year. Worldwide, more than 1.5 billion persons are infected with STH and more than 200 million are infected with schistosomiasis (Chitsulo, Engels, Montresor, & Savioli, 2000; Hotez, 2014; Lozano et al., 2012). It is well known that valleys and basins with lakes, rivers, and wetlands, with a warm and humid (tropical) climate, provide suitable conditions for various intestinal helminths, parasites and intermediate hosts (Currie, 2001; Degallier et al., 2010). Transmission risk increases when population groups engage in activities with direct exposure to intestinal parasites e.g. subsistence agriculture involving digging with hoes, working with bare hands and walking barefooted (Mote, Makanga, & Kisakye, 2005).

Thus, the spatial pattern of these infectious diseases is linked both to the biophysical and socio-economic environment and influenced by the agro-

ecosystem of wetlands (Gatrell & Loytonen, 1998). In this perspective, changes in the use of wetlands can represent a risk for emerging and re-emerging diseases. In short, environmental change (ecologic, socio-economic, loss of biodiversity, climate change), can have an impact on disease transmission (Eisenberg et al., 2007; Ramsar, 1971) as summarized in Figure 3.

Health records show that soil-transmitted helminths are widely distributed in most African countries, endemicity is facilitated *i.e.* by a number of climatic factors, including high temperature, and heavy precipitation. These two factors account for high humidity, the single most important parameter that favors the hatchability of geohelminth ova in the soil (WHO, 2003). Similarly, schistosomiasis is endemic in many countries (Sacko et al., 2011) and the second most important parasitic disease after malaria in Africa (Bomblies, Duchemin, & Eltahir, 2008; Mutuku et al., 2009; Vanek et al., 2006).

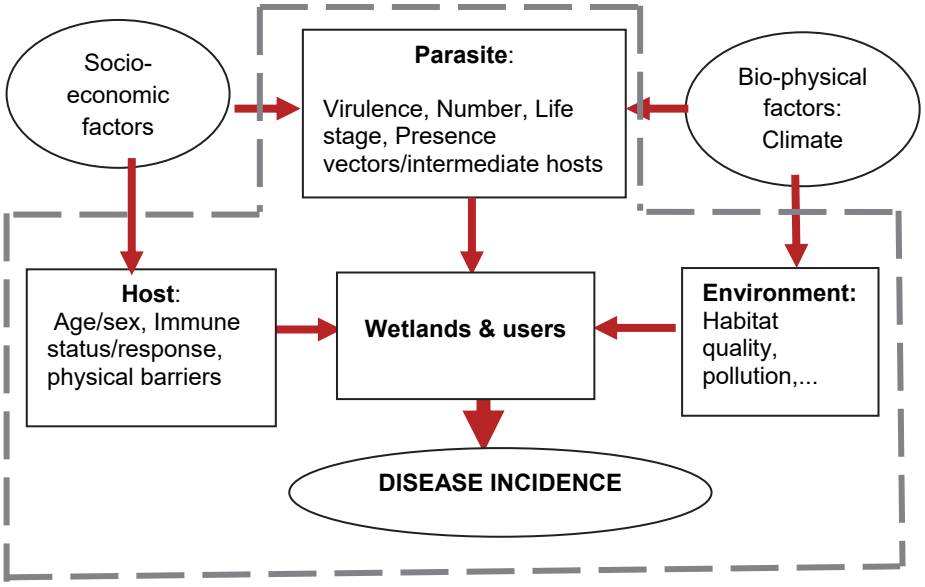

Figure 3: A simplified diagram of interconnected factors for Neglected Tropical Disease NTD monitoring.

In addition to suitable climatic conditions, wetland reclamation, especially for rice cropping, increases infection risk of local communities using traditional cropping systems with living conditions characterised by poor levels of hygiene and substandard sanitation.
In East African countries people living in rural areas are commonly infected with STH (Mupfasoni et al., 2009). In Rwanda, after malaria, schistosomiasis and STHs are a significant public health problem (TRAC+, 2008), as illustrated in Figure 4.

1.2 Why understanding the association between wetland environment and the spatial and temporal distribution of schistosomiasis and STHs?

Rwandan wetlands are rapidly converted from natural reserves with rich biodiversity to agricultural wetlands due to population pressure, which has increased demand for food production. Therefore, environmental laws and policies, advocating for a 'wetlands ecosystem health and human health', have been elaborated during the last decade. Unfortunately, their effective implementation is missing important baseline information for accurate spatial delineation and characterisation of wetlands, which is a pre-requisite for effective planning and sustainable use of wetlands.

Climate variability, wetland change and disease pattern assessment
In the Rwandan context, 'wetland' is defined as all lowlands and comprises the entire valley bottom, both the well-drained and wet areas (REMA, 2009). Rwandan wetlands were originally delineated using low-resolution satellite images and other data with varying spatial and temporal resolution. The inaccuracy of this approach is hindering the national process of systematic land registration and titling for individuals having land adjacent to wetlands which are formally registered as public land. Use of appropriate geo-information tools, techniques, and earth observation data, an accurate and high-resolution DEM and topographic derivatives, and selected biophysical variables can enable accurate demarcation and appropriate characterization of wetlands. A spatial model combining topographic, hydrologic, pedologic and climatic variables was inexistent. With an accurate wetlands characterization, informed decisions can be made regarding appropriate use (to be protected? to be irrigated or drained? etc.) of wetlands.

Despite that, numerous water development projects are ongoing for intensifying agriculture in most Rwandan wetlands. A good example is rice cropping, which often calls for considerable engineering and other technical interventions to provide extra water, with new drainage networks for evacuating excess water during flooding seasons, fertilizers and organic manure, etc. The "created rice cropping landscape" is being accused, among other negative impacts, to lead to increased incidence of some soil and water-transmitted diseases (Degallier et al., 2010; Gu & Novak, 2005; Mutuku et al., 2009). This relationship, however, has not yet been investigated in Rwanda.

A limited number of Rwandan researchers acknowledge the role of environmental factors related to wetlands and the proliferation of soil and

water-transmitted diseases (Karuranga, 2011). Spatial modeling of intestinal helminths disease incidence is a prerequisite for risk assessment. This can be achieved by combining empirical methods of data exploration and models drawn from the statistical and epidemiological sciences, with spatial analysis tools relating them to environmental and socio-economic factors (Gatrell & Loytonen, 1998). Use of Geographical Information Systems (GIS), and Remote Sensing tools and techniques can contribute to better understanding of infectious disease distribution patterns (Nuvolone et al., 2011; Thomas, Fischer, Fleischmann, Bittner, & Beierkuhnlein, 2011). According to our knowledge, no single study has assessed the spatial pattern of schistosomiasis and STHs and analysed their association with climatic and other potential risk factors. Consequently, little is known about risk factors in Rwanda for schistosomiasis and STH transmission, especially the possible relationship with intensified agricultural use of wetlands, rice cropping in particular.

Disease data acquisition and spatial pattern detection approach

As illustrated in Figure 4, like malaria, intestinal parasites represent a significant public health impact in Rwanda. More than 75% of children are infected with intestinal worms (TRAC+, 2008). Children heavily infected are often physically and intellectually compromised by anemia, leading to attention deficits, learning disabilities, school absenteeism, and higher drop-out rates.

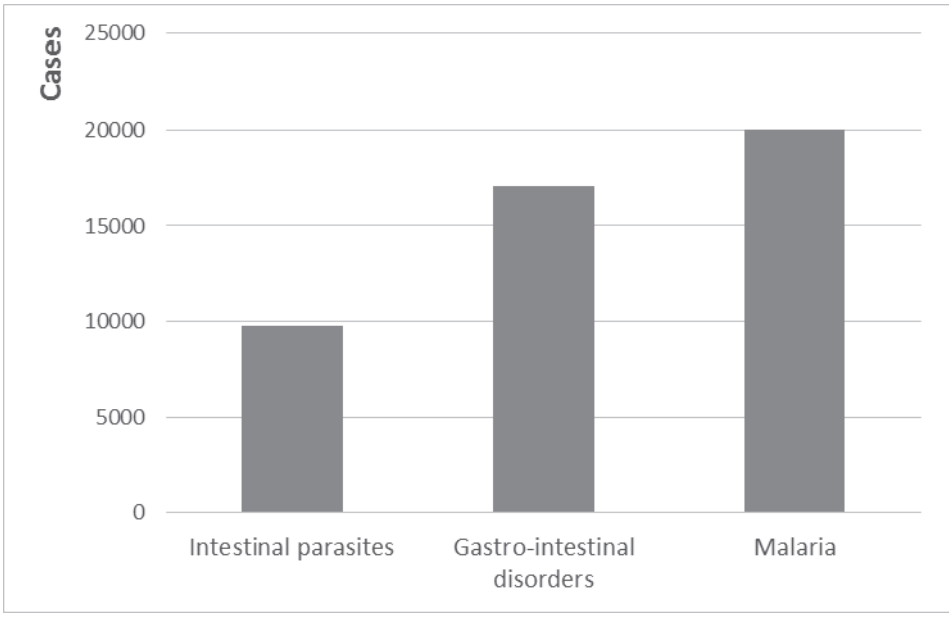

Figure 4: Outpatient visits to district hospitals in 2008. Source: Rwanda statistical yearbook, 2009

The disease statistics are reported in the national statistical yearbook of Rwanda (NISR, 2010b), but these are not used for disease mapping and modelling. The only existing nationwide schistosomiasis and STHs mapping was based upon a school-based prevalence survey conducted in 2007/2008.
The demand for statistics that accurately track health status, evaluate the impact of health programs and policies, and increase accountability at country and global levels is the concern of most health agencies since last decade (Chan et al., 2010). Over the last decade, the Rwandan Health Management Information System (RHMIS) has been designed and implemented in collaboration with the World Health Organisation (WHO)and several other partners.

Despite, the availability of systematically collected and spatially structured health data and advanced geostatistical tools and techniques for disease pattern detection (Brooker et al., 2009), existing studies on schistosomiasis and STHs are often based upon location specific prevalence surveys whereby survey outcomes are generally extrapolated to a larger spatial scale. A comparison between routinely collected health data and cross-sectional prevalence survey data should be done to assess if routinely collected health records compiled in the RHMIS have added or complementary value for guiding public health interventions.

Thus, an approach needs to be developed that can accurately identify and visualise the spatial and temporal pattern of schistosomiasis and STH transmission at a fine spatial resolution. Statistical and spatial analysis, furthermore, is combined to identify environmental and socio-economic risk factors associated with schistosomiasis and STH transmission. Finally, geostatistical modelling is done to forecast how future environmental change might impact on schistosomiasis transmission. Such outputs can help the Neglected Tropical Disease (NTD) control program to guide future prevention, control and elimination interventions. It can also contribute to better understanding of the linkages between agricultural development on the one hand and potential impact on disease transmission on the other.

1.3 Objectives of the thesis

The overall objective of this thesis was to assess the association between wetland environment and conversion and STHs and schistosomiasis in Rwanda. This research hypothesized that "wetland conversion into intensive agricultural use leads to increased intestinal helminths infections”. The specific objectives were:

1. To characterize Rwanda’s wetlands in a spatially explicit way that also allows a climate sensitivity assessment.

2. To determine the value of schistosomiasis incidence data, next to existing prevalence data, in detecting a spatial pattern of the disease. Given the fine-grained spatial resolution of health facility service areas in Rwanda, investigate if this could allow detecting focalized hotspots of the disease and its associated risk factors.
3. To investigate whether incidence data can be also used to detect the spatial distribution of STHs, and to assesses how the spatial variation of STH pattern is related to environmental risk factors.
4. To explore the spatiotemporal distribution dynamics of schistosomiasis clusters in Rwanda between 2001 and 2012, and to look at NTD control program impact.
5. To model the dynamic spatial pattern of schistosomiasis risk in Rwanda using a geostatistical model accounting for false zero cases, and to develop a scenario for identification of potential future risk areas

1.4 Description of study area and used data

The general motivation of this study is to recognize the spatial, temporal and spatiotemporal connections between wetlands, agricultural development projects and schistosomiasis and intestinal helminths incidence patterns. The analysis is mostly based on existing data analyzed using a combination of geo-information, exploratory techniques and spatial modelling tools.

Study area

Rwanda is a landlocked Country in Central – East Africa between 1° 04′ and 2° 51′ latitude South and 28°50′ and 30°50′ longitude East. Its surface area is 26.338 km^2 with a population of about 10 million (NISR, 2010a). Rwanda is administratively divided into four provinces and Kigali City, 30 districts, 416 sectors and 2148 cells (MINALOC, 2008). The topography is very irregular, offering a landscape of mountains, hills, and valley, with an altitude varying from 900 to 4.500 meter. The climate is moderated by the altitude with an annual average temperature of 19°C and an annual cycle of four seasons: short rainy season - short dry season - long rainy season - long dry season (Prioul & Sirven, 1981). The agriculture sector employs 90% of the active population and contributes to 41% of GDP and more than 72% of all exports (REMA, 2009). The land available on average for a household for agricultural use is below 1 ha. The development of wetlands and valleys for agricultural use is a response to population pressure. Wetlands contain large water reserves, have lower erosion risks and high natural fertility (Malesu et al., 2010).

Over time, Rwanda's health sector has experienced a profound evolution from traditional healing methods to faith-based health care during the colonial period, then centralized, and free provision of public health services

up to the early 1990's, to the current decentralized health care delivery system. This decentralization reinforces community participation in the management and financing of health services. The current public health care delivery system is a hierarchically organized three-tier system providing primary, secondary and tertiary health care. Primary health care is provided at sector level via one or more health facilities; secondary care is provided in the districts by a district hospital, and referral hospitals provide tertiary care at the regional scale. Each health facility reports to the district health office, which is responsible for the health facilities and services provided to the population of the district (see Figure 5). Community Health Workers (CHWs) function as an effective link between communities and public health care initiatives (Health Policy of Ministry of Health elaborated in 2005).

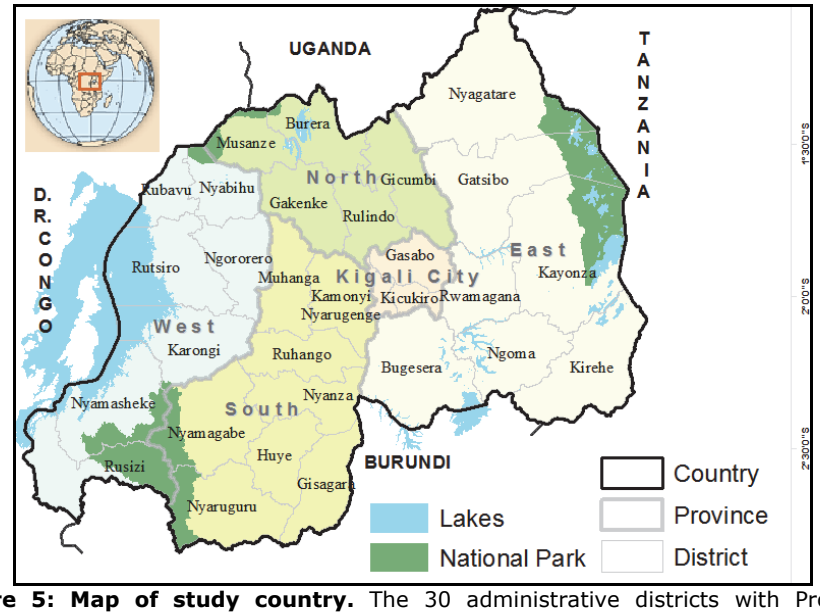

Figure 5: Map of study country. The 30 administrative districts with Province boundary of Rwanda shown in different colors.

Disease data and risk factor data

The number of confirmed cases of STHs and *schistosomiasis* for the years 2001 - 2012 were provided for this study by the Rwanda Biomedical Centre - Malaria and others parasitic diseases department (RBC/M&OPD). The same department also provided point location and prevalence data of the surveyed schools for STHs and schistosomiasis nationwide prevalence mapping in 2007/2008. The quality of systematically recorded cases at primary health facility level, is sound for several reasons. Firstly, the infection is diagnosed via Microscopic identification of eggs in stool samples in the laboratory of the health facility. Secondly, accessibility is unproblematic as patients can reach a primary health facility (health Centre, health post or Dispensary) within a walking distance of at most 5 km. Thirdly, Community-Based Health

Insurance (CBHI) makes that appropriate health care is affordable for everyone (patients pay only 10 % of the total cost of service and medication). Fourthly, community health workers actively stimulate patients to visit the health facility in case of suspected health problems. Fifthly, the behavior for health care seeking at the health facility has increased for most patients with symptoms of serious infection. Sixthly, the Rwandan Health Management Information System (RHMIS), routinely and systematically collects health records from individual health facilities using a web-based software platform (DHIS2) whereby each health facility enters their monthly health records directly into the national database (USAID/Rwanda, 2006).
Demographic data were extracted from the 2002 and 2012 Population, and Housing Census published by the National Institute of Statistics of Rwanda (NISR) and for the years in between the estimated population number and distribution were extracted from estimation and projections also reported by the National Institute of Statistics of Rwanda (NISR). The socio-economic factors such as school attendance levels, access to improved water sources for domestic use, proper sanitation at district and sector level were available from general census or integrated household living conditions surveys (EICV 2 & EICV 3) reported by NISR. In addition to the social and economic factors, biophysical factors need to be considered as potential risk factors at the highest spatial resolution. These data were acquired, from different governmental and non-governmental institutions and pre-processed in order to generate risk factors raster maps.

1.5 Outline of the thesis

This thesis consists of seven chapters successively structured as steps followed to achieve the specific research objectives. ***Chapter one*** gives a general introduction to the study. It introduces the scientific background, states the research problem, provides the research objectives and the description of the study area and datasets used.

Chapter two characterizes the wetland environment in Rwanda by combining terrain derivatives and climatic factors. In this chapter, the role of climatic factors in determining wetland occurrence probability is quantified using logistic regression. Also, the sensitivity of Rwandan wetlands to climatic factors at national and regional level is demonstrated.

The assessment of the capacity of confirmed cases of schistosomiasis in detecting the diseases spatial pattern is presented in ***Chapter three***. In this chapter, the known focalized disease distribution is visualized on the basis of routine health data collected at primary health facility level. This chapter also relates the observed schistosomiasis incidence spatial pattern to potential physical, ecological and climatic risk

In ***Chapter four*** the same approach is used for mapping soil-transmitted helminths cases. The spatial autocorrelation and sensitivity to different spatial scales of individual or combined STH distribution are clearly demonstrated. Also, the relationships between spatial clustering of intestinal helminths and the associated risk factors are explored.

The usefulness of routinely collected incidence data for identification of spatial patterns of schistosomiasis opens the door for spatiotemporal analysis in ***Chapter five***. The spatiotemporal clusters of schistosomiasis in Rwanda are presented here. The chapter furthermore, elaborates how the impact of the NTD national control program can help explain the localized clusters that were either persistent, emerging or disappearing over time.

In ***Chapter six***, a zero-inflated Poisson (Wu, Liu, & Davis, 2005) likelihood to fit a Bayesian Hierarchical spatiotemporal model was generated. This advanced modeling approach was selected due to the highly focalized and dynamic distribution of schistosomiasis clusters in time and space with many HFSA's having (true and false) zero cases. Also, the potential future distribution of the disease risks areas was explored.

To conclude, in ***Chapter seven*** a synthesis is provided of the findings from the descriptive and modelling approaches used. Possible implications of wetland conversion for intensive agricultural use on the NTD control program are also discussed.

Chapter 2: Regional climate sensitivity of wetland environments in Rwanda: the need for a location-specific approach[1]

[1] This chapter is based on a published paper: Nyandwi, E., Veldkamp, T., & Amer, S. (2016). Regional climate sensitivity of wetland environments in Rwanda: the need for a location-specific approach. Regional Environmental Change, 16(6), 1635-1647. doi:10.1007/s10113-015-0905-z.

Abstract

Wetlands are sustaining large communities of people in Rwanda where 10 % of its surface consists of many local wetlands. Sustainable future management of these numerous wetlands requires a reliable inventory of their location and a dynamic quantitative characterization that allows assessment of their climate change sensitivity. The aim of this study was to assess the importance of climatic factors in determining wetland location at different regional scales. Wetland locations were analyzed and statistically modeled using their location factors with logistic regression. Wetland location probability was determined using topographic (elevation, slope), hydrological (contributing area) and climatic (temperature and rainfall) location factors. A wetland location probability map was made that demonstrated a calibration accuracy of 87.9 % correct at national level compared to an existing inventory, displaying even better fits at the sub-national level (reaching up to 98% correct). A validation accuracy of 86.2 % was obtained using an independently collected dataset. A sensitivity analysis was applied to the threshold values used as a cut-off value between wetland/non-wetland, demonstrating a robust performance. The developed models were used in a sensitivity scenario analysis to assess future wetland location probability to changes in temperature and rainfall. In particular, wetlands in the central regions of Rwanda demonstrate a high sensitivity to changes in temperature (1% increase causes a net probable wetland area declined by 12.4%) and rainfall (+ 1% causes a net increase by 1.6%). This potentially significant impact on wetland number and location probability indicates that climate-sensitive future planning of wetland use is required in Rwanda.

Keywords: Wetland management; Topography; Climate change; Spatial scales, Probability model.

2.1 Introduction

Wetland environments encompass the transitional zone between land and water where the land is covered by shallow water or with a water table at or near the surface (Cowardin, Carter, Golet, & LaRoe, 1979, p. 3). Alternatively, wetlands are defined as spatial units having ecosystems associated with the long-term inundation of the soil (Keddy, (2010). This latter description emphasizes the eco-hydrological and geomorphological characteristics of wetlands. Consequently, various researchers have developed wetland modeling approaches on the basis of hydrological and geomorphic characteristics (Albert, Wilcox, Ingram, & Thompson, 2005; Brinson, 1993; Large & Petts, 1996; Xie, Liu, Jones, Higer, & Telis, 2011). To a lesser extent wetlands have characterized on the basis of topographic, hydrological as well as climatic factors (Mendoza-Sanchez et al., 2013; Ralph & Hesse, 2010). Although these three components can be identified separately, it is apparent that there is a considerable interdependence between them. The link between water availability and local hydrodynamics becomes more enlightened when the contributing area (CA) or ecological factors are considered too (Curie, Gaillard, Ducharne, & Bendjoudi, 2007; Zhou, Gong, & Liu, 2008).

In Rwanda, wetlands are recognized as an important natural resource and explicitly considered in national planning and policy. In 2008 the Rwanda Environmental Management Authority conducted a national inventory and mapping of all wetlands, lakes, and rivers (REMA, 2008). The delineation and classification accuracy of the inventory, however, were quickly questioned as numerous inconsistencies were identified during the nation-wide land registration process that started two years later. These inconsistencies point to the fact that wetland bodies are changeable over time. During 'wetter' years more wetland will exists than during 'drier' years. Human interventions such as artificial drainage or dam constructions will also affect wetland locations and size. A reliable demarcation of wetland in Rwanda is needed since rapid conversions change a lot on biophysical and hydrological conditions (Hansen, Brown, Dennison, Graves, & Bricklemyer, 2009). Given the dynamic properties of wetlands, a probability approach instead of a static delineation (mapping) approach is considered more appropriate and flexible for future sustainable development.

Up to now, modeling approaches to characterize wetlands are generally based upon slope (wetlands tend to have flat slopes) and water accumulation derived on the basis of contributing area (Buis & Veldkamp, 2008b). Digital Elevation Models (DEM) are readily accessible nowadays and automated procedures exist to compute terrain derivatives (slope, curvature, topographic index, variance in slope, etc.) which can discern real depression

features from spurious ones (Hogg & Todd, 2007; Temme, Schoorl, & Veldkamp, 2006; Yamazaki et al., 2012).

Less commonly used for wetland modeling are the more dynamic climate related data that determine water availability and the local net water balance. Net water availability is a function of rainfall and evapotranspiration in the wetland catchment and the wetland itself. Evapotranspiration, in turn, is a function of vegetation, temperature, and humidity.

In short, improved modeling of wetlands and their spatial and temporal dynamics requires the use of topographic, hydrologic and climate variables (Adam, Mutanga, & Rugege, 2010). An essential advantage of such a modeling approach lies in its capability to determine how climate changes can affect wetlands in a spatially disaggregated manner.

It is, therefore, the objective of this paper to characterize Rwanda's wetlands in a dynamic spatially explicit way that also allows a climate sensitivity assessment. As argued above a probability approach is followed in order to allow quantification of wetland probability in space and time. Previous investigations have demonstrated that the spatial scale (grain and extent) of analysis and management (Fóti et al., 2014; Kok & Veldkamp, 2011; Veldkamp & Fresco, 1997) can affect outcomes of land system properties.

2.2 Materials and Methods

Study area

Rwanda is a country with considerable bio-geophysical diversity, with the main gradient from the lower and drier east to higher and wetter western part of the country. Rwanda, as a case study, offers the opportunity to identify the underlying factors that drive the formation of wetlands at different spatial scales.

Rwanda is located on the great East African plateau, which includes the continental water divide between the Nile and Congo rivers. Its climate is moderated by the altitude with an annual average temperature of 19°C and an annual cycle of four seasons: short rainy season, short dry season, long rainy season, long dry season (Prioul & Sirven, 1981). The Highland passes in a north–south direction through the western part of the country. To the west of the divide, the land drops abruptly to Lake Kivu in the Great Rift Valley. To the east, elevation gradually declines from the central plateau to the lowland of the eastern border of the country (Figure 6), with the following natural units:

1) Ten agro-ecological zones (Paez & Scott): Buberuka Highlands (BH), Bugarama Plain (BP), Central Plateau (CP), Congo Nile Watershed Divide

(CNWD), Cyangugu Backside (CB), Eastern Ridge and Plateau (ERP), Eastern Savanna (ES), Kivu Lakeside (KS), Mayaga Plateau and Central Bugesera (Rochlin, Turbow, Gomez, Ninivaggi, & Campbell) and Volcanic Summits and High Plains (VHP).

2) Two main major drainage basins: the Nile basin in the east covering 67% of the country and the Congo basin in the west which covers the remaining 33% (REMA, 2008). The two drainage basins are subdivided into ten watersheds: Akagera (NAKA), Akanyaru (Kaliappan et al.), Kivu (CKIV), Mukungwa (NMUK), Mulindi (NMUL) Muvumba (NMUV), Mwogo (NMWO), Nyabarongo amont (NNYT), Nyabarongo aval (NNYL) and Rusizi (CRUS).

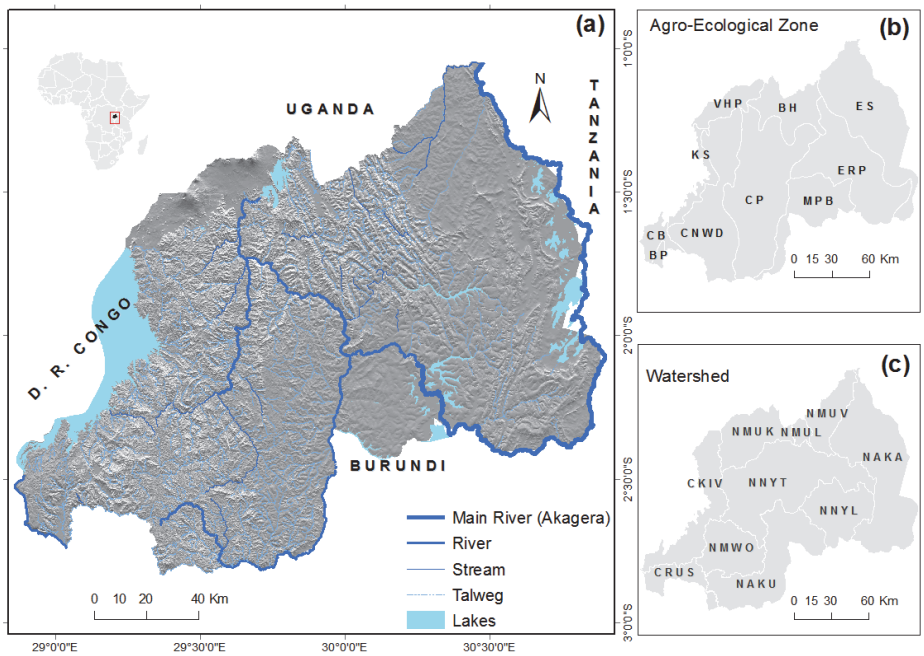

Figure 6: Inland lakes and a hydrological network of Rwanda. With Akagera as the main river, collecting water from 67% of the country into the Nile river basin (**a**). The delimitation of regions used in sub-national modeling is visualized by two maps in the right corner, namely ten Agro-ecological zones (**b**) and ten main watersheds (**c**).

Potential geophysical location factors for wetlands in Rwanda

Topographic factors

The Rwanda National Land Use and Development Master Plan Project produced a high-quality DEM with a 10 x 10-meter raster cell size for Rwanda (Swedesurvey, 2010). From the DEM a point grid was extracted with an interval of 10 meters. The DEM was provided by the Rwanda Natural Resources Authority for this study.

In addition to elevation data, the DEM was used to derive the slope gradient which is used as a landform characteristic in this study. The slope gradient data set, which indicates the rate of maximum change in elevation for each cell of the DEM, was calculated using the following equation:

$$\text{Slope (\%)} = 100.\text{Sqrt}\left(\left(dx^2 + dy^2\right) / \left(2^* \text{ grid size}\right)^2\right) \qquad \textbf{\textit{(Eq. 1)}}$$

Where ***dx*** is the height difference (m) of the pixel in x-direction and ***dy*** is the height difference (m) of the pixel in ***y***-direction and grid size is the length of the side of one grid that assume the length of each side is the same.

A second topographic derivative that was extracted from the DEM is the Contributing Area (CA)**,** which represents a hydrological characteristic. The CA dataset was generated using LandscApe ProcesS modeling at mUlti-dimensions and Scales (LAPSUS) techniques, as this resolves the issue of divergent flows and multiple depressions (Temme et al., 2006).

Climate factors

A network of 183 meteorological stations (see Figure 7a) distributed throughout the country records rainfall, temperature, evapotranspiration and relative humidity. Annual averages of these data were measured for most stations using 60 years records starting from 1950 to 2010 and obtained from the Rwanda Meteorological Center. We interpolated these data using the thin-plate smoothing spline algorithm as proposed by Hijmans et al. (2005). The produced maps were transferred to raster datasets with a 10 meters cell size following the approach of Mukashema et al.,(2014).

Wetland, Agro-ecological zone, and Watershed data

In addition to above-described factors data, several additional GIS and remotely sensed datasets were used. For wetland-upland sampling, wetlands, lakes and country boundary data were used. The wetland data used originates from the inventory conducted in 2008 by REMA, which produced 860 wetlands covering a total surface of 278,536 ha, corresponding to 10.6% of the total country area. The administrative boundary of Rwanda, agro-ecological zone boundaries (defined on the basis of altitude, rainfall and soil characteristics) and main watershed boundaries were obtained from the Centre of GIS and Remote Sensing of the University of Rwanda. In addition, field data were collected at 299 locations (see Figure 7b), randomly spread over the country during March-July 2013 and December – April 2014. For each geographical location the X, Y coordinates of site categorized as wetland/upland were recorded. The empirical data generated from field work are used for independent validation of the statistical models.

Data pre-processing and processing

Data preparation

In this study, we limit ourselves as much as possible to the use of original measured data. The rationale for this is to avoid data redundancy, between derivatives that share the same data source. A terrain derivative such as the topographic index, which is frequently used in simpler models that do not consider hydrological data (Chirico, Western, Grayson, & Blöschl, 2005) is therefore not included in the analysis. Evapotranspiration and relative humidity are highly correlated with (often calculated using) temperature and rainfall and therefore also excluded from the analysis (Creed, Sanford, Beall, Molot, & Dillon, 2003; Hogg & Todd, 2007). The most important and independent topographic and climatic factors to predict the occurrence of wetlands in Rwanda are summarized in Table 1

To correspond with the spatial resolution of the DEM, all factor raster data were resampled a 10 meters cell size. Another motive for choosing a high spatial resolution is to also capture smaller-sized wetlands of which there are many in Rwanda. Factor data were subsequently standardized into a range from 0 to 1 using a minimum-maximum linear transformation. This normalization is necessary to ensure comparison of values of the same range, which is an important prerequisite for multiple variable analysis with continuous variables *(Long, Nelson, & Wulder, 2010).*

Table 1. The most original factors used for wetland occurrence modeling

Category	Factors	Description
	Elevation	Preprocessed high-resolution DEM of 10 m cell size produced using ortho-photographs, 2008
	Slope	Slope calculated using DEM
	Contributing area	Hydrological modeling – Multiple flow accumulation using LAPSUS Model
	Rainfall	Total rainfall in mm measured at meteorological station, interpolated to obtain values for the whole study area
	Temperature	The annual mean measured at meteorological station interpolated to obtain values for the whole study area

Sampling design

Training sample points were randomly selected from wetland and non-wetland locations. To prevent spatial auto-correlation data points were sampled from the REMA map of wetland for training (50%) and validation (20%) purposes, respectively. Wetland polygons were merged to allow random distribution of points to be generated from total coverage. The non-wetland area was identified using ArcGIS overlay functionality by excluding

all wetlands and water bodies. The sampled point dataset with 1 single attribute (0 = non-wetland, 1 = wetland) was further populated by extracting the associated elevation, slope, contributing area, rainfall and temperature values from the respective factor raster data. The fully populated attribute table was then exported and analyzed using SPSS Statistics 20.
The same process was repeated for each agro-ecological zone and for each watershed.

Logistic regression analysis (LRM)

Given that our dependent variable is dichotomous (either wetland or no wetland) and we have a set of predictor variables, a binary logistic regression approach is a suitable statistical technique for our analysis (Burns & Burns, 2009). Binary logistic regression determines the impact of multiple independent predictor variables presented simultaneously to predict membership of one or other of the two dependent variable categories (Ravi & Kulasekaran, 2013). The typical link purpose is logit function; which in this case related the logarithm of the odds ratio of expected value of wetland probability to linear covariant of elevation (Elev), slope (Slp), collecting area (CA), temperature (Temp) and rainfall (Rain). If the [π(x)] represent the expected value of probability of wetland occurrence at location *i,* then the logistic regression model is [Equation 2]:

$$\text{Logit}\,[\pi(x)_i] = \beta_0 + \beta_1 \times \text{Elev}_i + \beta_2 \times \text{Slp}_i + \beta_3 \times \text{CA}_i + \beta_4 \times \text{Temp}_i + \beta_5 \times \text{Rain}_I \qquad \textbf{(Eq. 2)}$$

Where x is the probability of wetland occurrence; β_0, β_1, β_2, β_3, β_4 and β_5 are the regression coefficients to be estimated for the independent variables.

The T-Wald statistic (*z*-value) and associated probabilities provide an index of the significance of each predictor in the equation. The Wald statistics asymptotically follow chi-square (χ^2) distribution (Kharkar & Bowalekar, 2014). Also, the Exp(*B*) (odds ratio) presents the extent to which raising the corresponding measure by one unit influences the probability of the dependent variable (Burns & Burns, 2009), wetland occurrence in this case. In addition, the Nagelkerke's R^2, meanings the power of explanation of the model, was also considered.

Creation of wetland probability geo-database

The wetland probability geodatabase was computed (with ArcGIS 10.1) using the coefficients of the fitted regression model for the significant predictor variables. The resulting raster map displays the probability of wetland occurrence, ranging from low probability (yellow) to high probability (dark blue).

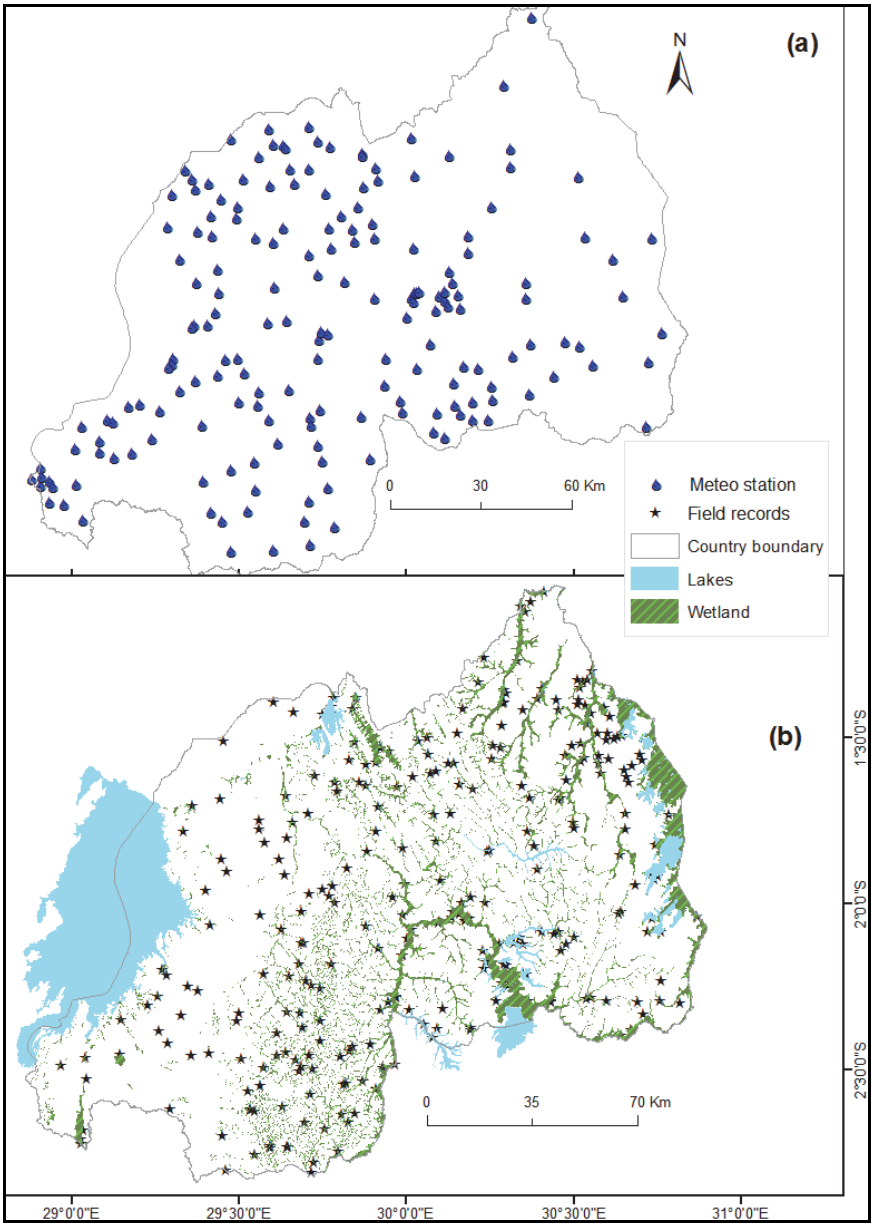

Figure 7: Location of (a) 183 Rwandan (agro) meteorological station and (b) Field dataset sampling sites. The sampling sites are overlapped to Rwandan Wetlands as defined by the inventory and mapping of wetland by REMA (2008).

Sensitivity of the LR model to spatial scale and climate factors

As discussed above, logistic regression analysis was done to assess whether topographic and climatic factors can reliably determine the probability of wetland occurrence. Because of the potential scale sensitivity, it was relevant

to explicitly assess the influence of spatial scale on the studied system properties. In order to adequately capture such potential scale effects, the wetland characterization of Rwanda was done at two different spatial scales. In the first instance, the analysis was done at the national level for Rwanda as a whole. After the national extent, the analysis was repeated at subnational level using the same predictors but now in a spatially disaggregated manner for each individual agro-ecological zone and watershed, respectively.

Differences in model performance at national and subnational scale were compared to assess whether model fits are indeed sensitive to a spatial extent. The subnational model results were subsequently reassembled to produce a new composite probability geodatabase for the country as a whole.

To illustrate how climate change can impact on spatial wetland occurrence, a scenario with an overall increase of 1% for temperature and rainfall respectively, were calculated with the developed logic regression models. The slight increase of 1 % was selected referring to the trend analysis of the annual mean temperature in Rwanda during last 52 years (Safari, 2012). New wetland probability maps were calculated and compared. In the two scenario's cells, a probability value used as the threshold was defined after the sensitivity analysis of the probability map of wetland occurrence produced using logistic regression models. Different thresholds for different scale and models were expected.

Model performance and validation

In addition to statistical significance testing, the performance of the model was independently assessed using the field-based data set, collected at 299 randomly selected wetland and upland sites. The location of a ground-based dataset of wetland/upland as points map is provided with Figure 7b.

In this case, we also made a sensitivity analysis for testing the robustness of the results of our models at national and regional level using XY Scatter plot of different probability value from 0.6 to 1.

2.3 Results

Predictors of wetland occurrence at national level

A test of the full model against a constant only model is statistically significant, indicating that the predictors together can significantly distinguish between wetland and non-wetland locations (χ^2= 547.271, p <.000 with *df* =5). Also, the classification error rate changes from 50% (Step 0); by adding the variables, we can now predict with 85.0% overall accuracy (87.9% for wetland and 82.1% for non-wetland). Nagelkerke's R^2 of 0.628.

The national model has elevation, slope, rainfall, temperature and contributing area as significant independent factors considering the Wald's test. The strong contribution of single variable expressed by the Exp (*B*) coefficient is attributed to both topographic (slope) and climatic (temperature) factors as summarized in Table 2.

Table 2. Predicting factors of wetlands occurrence and their statistical coefficients

Parameters	B	S.E.	Wald	Exp(B)
Elevation	-0.008	0.001	65.136	0.992
Slope	-0.213	0.020	108.038	0.808
Contributing area	0.001	0.001	28.060	1.000
Rainfall	0.002	0.001	11.475	1.002
Temperature	-0.815	0.159	42.096	**0.357**
Constant	30.650	4.395	48.625	2.047E+13

Spatial scale dependence

Agro-ecological zone level

The analysis of agro-ecological zone level results in 10 logistic regression models, each with its own coefficients. Table 3a summarizes the significant drivers of wetland occurrence for each agro-ecological zone.

For Mayaga Plateau and Central Bugesera (Rochlin et al.) all five potential predictor variables are included in the fitted model. The odd ratios clearly indicate that slope and temperature are the most forcing factors predicting wetland occurrence.

In the agro-ecological zones of Buberuka Highlands (BH) and Bugarama Plain (BP), elevation, temperature, and slope are the main factors predicting wetlands. As can be seen in Figure 6b, BH is characterized by an altitude and therefore temperature change. In the Bugarama Plain, located in the southwestern and lowest part of the country, temperature and slope are the most important predictors.

For the Central Plateau (CP), with its dense drainage network, the main predicting factor is temperature. Slope and elevation have only minor influence; rainfall is not a significant factor.

The Congo Nile Watershed Divide (CNWD), Cyangugu backside (CB), Eastern Ridge and Plateau (ERP) and Kivu lakeside (KS) all have topographically sensitive wetlands. slope, elevation and contributing area are respectively the best predicting factors.

In the Eastern Savanna (ES) wetland occurrence is only and strongly associated with the slope.

The volcanic summits and high plains (VHP) model also represents an exception by not including 'elevation' but instead introducing 'rainfall' as a predictor for wetlands occurrence.

Temperature is a highly significant climate predictor in the Buberuka highland, Bugarama plain, Central plateau, Mayaga Plateau and Central Bugesera and Eastern Ridge and Plateaus. Rainfall predicts wetland occurrence with moderate significance in the volcanic summits and high plain zone.

Model fit improved with an increase of Nagelkerke's R^2 from 0.628 to 0.910 and the prediction success improved with overall accuracies ranging from 0.85 to 0.98 for seven out of ten AEZ. Model performance decreased in the Central plateau and the Cyangugu backside (from 0.628 to 0.524 with an overall accuracy of classification decreased to 81.7%). In general, an increase indicates that wetland that is better captured by the regional models than the model at the national level (Kok & Veldkamp, 2011).

Watershed level

At the level of watersheds, the regional models also performed differently in comparison to the national model, as summarized by Table 3b.

Wetlands in the Akagera watershed are predicted by elevation, slope, and rainfall with a highly significant effect of the slope. The Akanyaru watershed retains the same predictors as the national model, but with improved statistical performance.

For the Mukungwa catchment, elevation and temperature are not included in the model as explanatory factors. For the Mulindi and Rusizi watersheds, both temperature and contributing area are insignificant. The most significant factors for these watersheds are respectively slope, elevation and rainfall. In Muvumba watershed wetlands characteristics can be explained by elevation, slope, contributing area, and rainfall.

In Nyabarongo amont and Nyabarongo aval, elevation, slope, and temperature predict wetlands occurrence with temperature as the most contributing predictor.

Overall, both topographic and climatic components are equally important in predicting wetland occurrence at the watershed level, with a minimum of

three drivers per model. The rainfall was included in the model for most watersheds as a significant contributing factor.

Considering the Nagelkerke's R^2 and overall accuracy, three different outcomes can be noticed with respect to model outputs: (1) In some watersheds the model performance is more or less equal to the national level namely Akagera, Kivu, Nyabarongo amont and Mwogo catchments; (2) the subnational models performed not very well in the Muvumba and Akanyaru catchments; (3) in four cases (Mulindi, Nyabarongo aval, Mukungwa and Rusizi) the catchment model performed much better reaching 90% and higher.

A visual inspection of the wetland probability maps at national, agro-ecological zone and watershed level shown in Figure 8 reflects: (a) quite similar probability for national and watershed-based models in the northern and western part of the Country; (b) changing probabilities with all models in south and central part of the country and (c) a more wet eastern part of the country (i.e the Akagera watershed).

Table 3. Best predicting Logistic regression models outputs

a. At Agro-Ecological Zone level

Nr	Level	Variables	Coeff (b)	S.E	Wald (z-value)	Exp (B)	Nagelkerke R^2	Class. accuracy (%ge)
		Elevation	-.018	.005	10.438	.982		
		Slope	-.448	.131	11.680	.639		
		Tempera ture	-2.209	.979	5.094	.110		
		Constant	77.220	26.297	8.623	3439E+33		
		Elevation	-.066	.029	5.300	.936		
		Slope	-.513	.208	6.071	.599		
		Tempera ture	-4.854	2.378	4.166	.008		
		Constant	177.509	81.695	4.721	1.233E+077		
		Elevation	-.009	.002	32.274	.991		
		Slope	-.142	.019	56.555	.867		
		CA	.0003	.000	10.939	1.000		
		Tempera ture	-.826	.259	10.168	.438		
		Constant	31.483	6.803	21.418	4.71E+13		
		Elevation	-.004	.001	10.169	.996		
		Slope	-.183	.040	21.379	.833		
		CA	.0004	.000	3.585	1.000		
		Constant	10.290	2.894	12.641	29438.757		
		Elevation	-.009	.004	5.704	.991		
		Slope	-.083	.045	3.403	.920		
		CA	.001	.000	6.677	1.001		
		Constant	14.583	6.022	5.864	2153498.153		
		Elevation	-.031	.007	19.296	.969		
		Slope	-.346	.094	13.496	.707		
		CA	.006	.007	.662	1.006		
		Tempera ture	-2.480	.872	8.089	.084		
		Constant	96.175	25.742	13.959	5.864E+041		
		Slope	-1.023	.161	40.537	.359		
		Constant	2.275	.362	39.589	9.728		
		Elevation	-.008	.003	10.910	.992		

	Lakeside	Slope	-.200	.052	15.124	.818		
		CA	.001	.001	5.942	1.001		
		Constant	15.014	4.282	12.294	3315317.837		
		Elevation	-.036	.009	17.252	.965		
		Slope	-.452	.107	17.984	.636		
		Rainfall	.011	.005	5.312	1.011		
		Temperature	-3.284	.941	12.174	.037		
		CA	.0003	.000	2.848	1.000		
		Constant	107.184	29.188	13.485	3.544E+046		
		Slope	-.305	.142	4.585	.737		
		Rainfall	-.023	.008	8.988	.978		
		Constant	30.389	9.778	9.659	1.58E+13		

b. At Watershed level

S/N	Level	Variables	Coeff (b)	S.E.	Wald (z-value)	Exp (B)	Nagelkerke R^2	Class. accuracy (%ge)
		Elevation	-.015	.002	68.009	.985		
		Slope	-.623	.054	131.637	.536		
		Rainfall	.006	.001	21.666	1.006		
		Constant	16.819	2.032	68.505	20150295.722		
		Elevation	-.005	.002	8.964	.995		
		Slope	-.213	.037	33.182	.808		
		CA	.011	.008	1.900	1.011		
		Rainfall	.006	.001	17.509	1.006		
		Constant	2.229	1.685	1.750	9.290		
		Elevation	-3.135	75.889	.002	.043		
		Slope	-2.871	74.252	.001	.057		
		Rainfall	1.628	40.531	.002	5.094		
		Constant	3898.285	94974.672	.002	.000		
		Slope	-.177	.020	82.661	.837		
		CA	-.015	.010	2.009	.985		
		Rain	-.014	.003	30.120	.986		
		Constant	19.136	3.178	36.257	204468262.497		

	Elevation	-.002	.001	6.401	.998		
	Slope	-.118	.013	79.747	.888		
5 Kivu	Rain	.001	.000	3.960	1.001	0.547	80.6
	Temp	.616	.221	7.782	1.852		
	Constant	-8.323	5.557	2.243	.000		
	Elevation	-.002	.001	.998	.998		
	Slope	-.395	.013	.674	.888		
	Rain	-.004	.000	.996	1.001		
	Constant	12.461	.221	258094.774	1.852		
	Elevation	-.019	.002	108.585	.981		
	Slope	-.127	.015	73.558	.881		
	CA	.0003	.000	3.412	1.000		
	Rain	.007	.001	28.629	1.007		
	Temp	-2.633	.276	90.907	.072		
	Constant	74.463	7.894	88.978	2.18E+32		
	Elevation	-.048	.006	58.507	.953		
	Slope	-.315	.055	32.522	.730		
	Temp	-2.032	.441	21.222	.131		
	Constant	109.678	15.864	47.797	4.291E+047		
	Elevation	-.007	.001	74.256	.993		
	Slope	-.164	.013	167.528	.849		
	Temp	-1.069	.154	48.060	.343		
	Constant	33.666	4.111	67.068	4.18E+14		
	Elevation	-.007	.001	30.970	.993		
	Slope	-.109	.016	44.510	.897		
	Rain	.008	.002	16.408	1.008		
	Temp	1.434	.366	15.381	4.197		
	Constant	-24.345	8.643	7.933	.000		

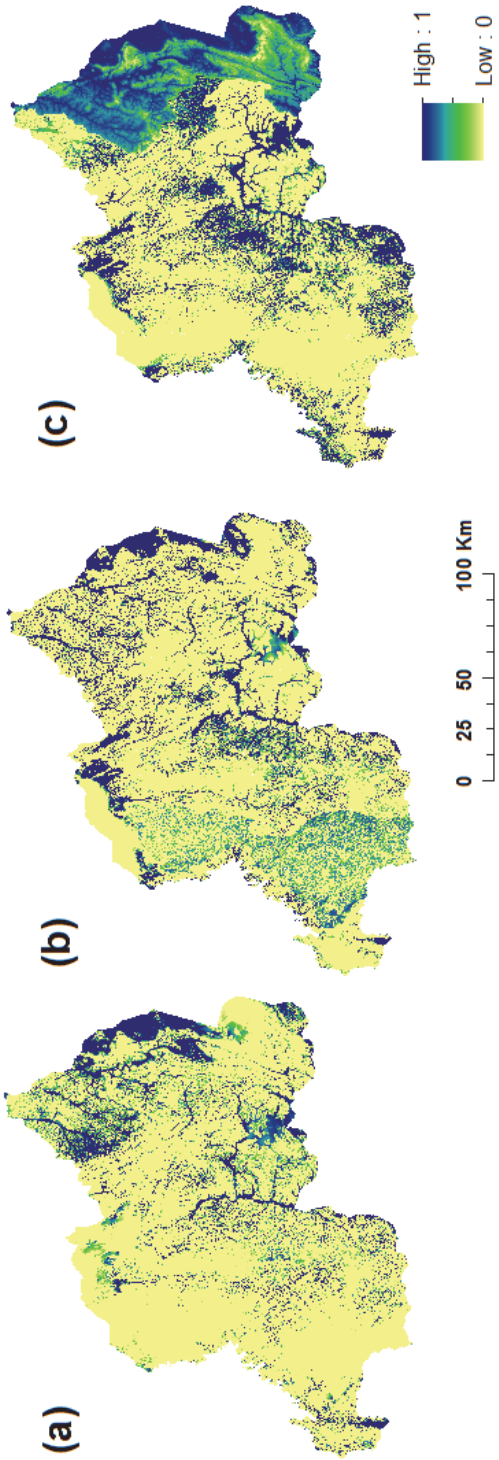

Figure 8: Wetland probability maps. Probability maps are generated using logistic regression model at national level (**a**), Composite probability map with AEZ models (**b**) and composite probability map with watershed models (**c**)

Exploring climate sensitivity of Rwandan wetlands.

The sensitivity scenario of a 1% increase in respectively rainfall and temperature results in considerable changes in the spatial pattern and dimension of wetland probabilities. An increase in rainfall is expected to lead to more/larger wetlands whereas an increase in temperature will lead to more evapotranspiration and therefore result in a reduction of wetlands. The accuracy of the threshold probability values was tested independently by a reliability analysis and will be discussed later.

Table 4 summarizes how a 1% increase in rainfall or temperature will affect the extent of wetlands. If we use the national model a 1% overall increase in rainfall will result in an overall growth of 1.6% of wetland area (using a threshold probability of 0.6). A 1% increase in temperature will have a considerably stronger impact and results in a nation-wide decline of 12.4% of probable wetland area. To determine the effects of climate change at a spatially disaggregate level Table 5 also presents a comparison at agro-ecological zone using Mayaga Plateau and Central Bugesera as an example. In this case, we see that a 1% increase in rainfall will result in an expansion of probable wetland are of 16.2%. A 1% increase in temperature in Mayaga Plateau and Central Bugesera will have a very substantial impact as it will reduce probable wetland area by as much as 37.9%.

Table 4. Effects of different climate change scenario on wetland area at national level and at the most sensitive agro-ecological zone

Climate factors variation	Upland (in ha)	Wetland (in ha)	Proportion (%ge)		Gain/Loss of wetland cover (%)
			Upland	Wetland	
Country model with current Climate factors value	2,122,929	234,372	90.1	9.9	
Country model with increased Rainfall (1%)	2,115,167	242,134	89.7	10.3	1.6
Country model with increased Temperature (1%)	2,174,541	182,760	92.2	7.8	-12.4
AEZ- MPB model with current Climate factors value	176,444	34,956	83.5	16.5	
AEZ- MPB model with increased Rainfall (1%)	162,946	48,453	77.1	22.9	16.2
AEZ- MPB model with increased Temperature (1%)	176,192	35,207	83.3	16.7	-37.9

More in general, the above examples demonstrate that the impacts of climate change will vary considerably from place to place. It also underlines the importance of using spatially disaggregate models as these are capable of identifying in which locations climate change will have more effect. The scenarios, also illustrated in Figure 9, clearly demonstrate the location specific regional sensitivity of climate change on wetland probability in Rwanda.

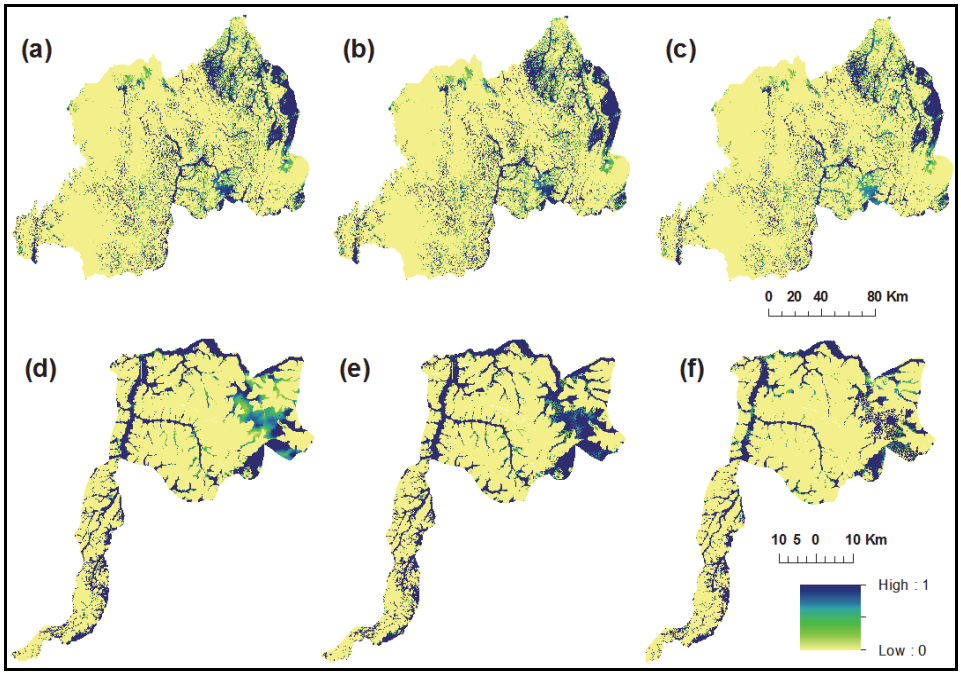

Figure 9: Wetland sensitivity to climate changes. At the national level: created with current climate factors value (**a**), with 1 % increase of rainfall (**b**) and with 1% increase of temperature (**c**). Sensitivity maps at the most sensitive Agro-ecological zone of MPB created using current climate factors value (**d**), with 1 % increase of rainfall (**e**), with 1% increase of temperature (**f**).

Wetland prediction models sensitivity analysis and accuracy assessment

A reliability analysis was complemented by using an independent dataset collected from the field. The results are summarized in Table 5. The topographic and climatic factors used in the best fitting statistical and spatial model explained 86.2 % of the probability of wetland conditions to occur. The number of field observations was not sufficient to conduct a similar validation exercise at each AEZ and watershed.

The sensitivity analysis of the used thresholds for wetland-non-wetland distinction was done by testing the robustness of the results of our models at national and regional level. The 85% of correct classification of wetland occurrence was reached with a probability value of 0.6. Wetland occurrence probability reaching 95% of accuracy/confidence as the target is reachable with a threshold of 0.6 of probability for the National model. The AEZ and watersheds, wetland occurrence probability map need to use mostly 0.7 and in few cases, 0.8 for probability maps with right visualization as illustrated by XY scatter plot in Figure 10. In general, it can be concluded that given the

steepness for the cumulative graphs in Figure 10., all thresholds are distinctive and robust in separating wetlands from non-wetlands.

Table 5. Accuracy of National wetland probability map using independent field dataset

Observed	**Predicted**		**Percentage Correct**
	Upland	**Wetland**	
Upland	83	21	79.8
Wetland	14	176	92.6
Overall Percentage			86.2

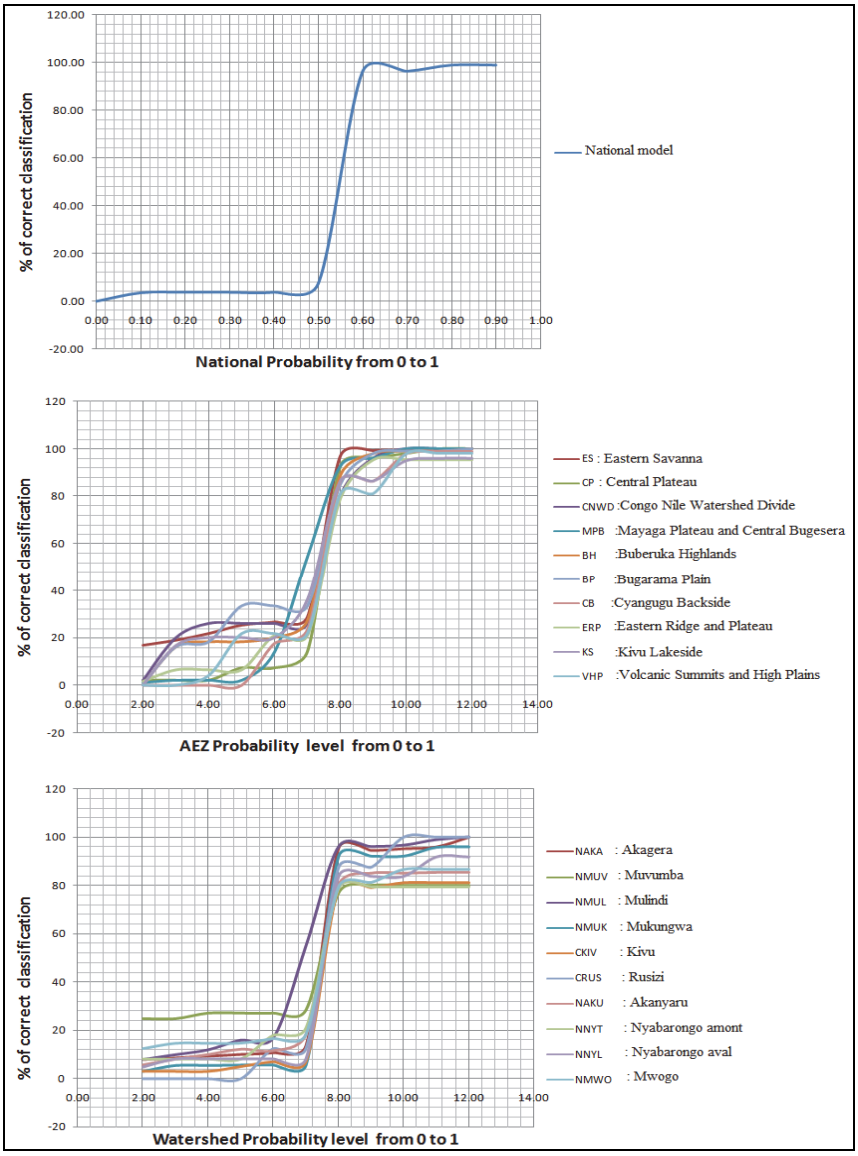

Figure 10: Sensitivity analysis using XY Scatter plots The scatter plots of probability value of wetland occurrence maps created using logistic regression models from national (**a**), AEZ to the regional level, respectively Agro-ecological zones (**b**) and Watershed (**c**).

2.4 Discussion

The use of topographic (elevation and slope), hydrological (contributing area) and climatic (rainfall and temperature) factors allowed the construction of statistically significant models for wetland characterization. The models give

satisfactory accuracy (86.22%) considering the highly complex nature of wetlands in Rwanda and the limited number of variables used.

The variables selected in the logistic regression models vary in number and significance (standardized betas coefficient and Exp (B)) at the different spatial scales used in the analysis presented. This indicates the highly regional specific nature of Rwandan wetlands. The potentially most dynamic wetland determining factors are climate related factors even though these are not significant for all regions. All regional models, however, contain the more static topographic factors. Clear differences in wetland characteristics and sensitivities exist between agro-ecological regions. When homogeneous and high spatial and temporal resolution climate data are available, the potential spatial dynamics of wetland in Rwanda can be better assessed. Such information, in turn, is highly valuable for their rational management. A one-size-fits-all management approach is clearly not going to be the most appropriate strategy for future policy development. We, therefore, recommend using a location-sensitive approach, as presented in our study to identify which wetland environments require extra attention in order to make them future proof.

Functional wetlands surfaces characterization drivers

Our study demonstrates that topographic and climatic factors are suitable to predict wetland occurrence in Rwanda at a level of 86.22 % accuracy. The importance of both factors was demonstrated at national and subnational levels.

Landforms are relatively stable landscape features which are dominant factors influencing the particular hydrological conditions (Hengl, Gruber, & Shrestha, 2003). Thus, slope and elevation appear to be consistent variables for the prediction of wetlands (Creed et al., 2003; Hogg & Todd, 2007).

Climatic components, one the other hand, are more dynamic and have varying importance as wetland occurrence predictor at the different scale levels. Climate variations or change is a prime driver of the water regime in the wetland ecosystem and can be a key factor for their characterization (Zhao, Stein, & Chen, 2011) and management. Several studies around the world revealed the importance of climate factors as a hydrological contribution at wetland and full catchment level (Mendoza-Sanchez et al., 2013; Nagumo, Sugai, & Kubo, 2013; Schulte, Veit, Burjachs, & Julià, 2009). The varying significance of temperature can be associated with its altitudinal degradation from bottom land to the hilltop (Dai & Huang, 2006; Melloh, Racine, Sprecher, Greeley, & Weyrick, 1999). For that reason, some wetland managers have already integrated climatic conditions such as rainfall and

temperature for assessing wetness conditions (Dimitriou & Zacharias, 2010; Gell, Mills, & Grundell, 2013).

Spatial scale sensitivity

The prediction of wetland occurrence probability at multiple spatial scales revealed variations of predicting factors and model performance.

In most cases, both climatic and topographic factors were combined in fitted models with different numbers of factors and different levels of significance. However, a few agro-ecological zones turned out to be exceptions as no single climatic component was included in the probability model. These are the Congo Nile Watershed Divide, the Cyangugu backside, the Kivu lakeside and the Eastern Savanna (Figure 6b & c). Out analysis suggests that in these regions the wetlands are less climate-sensitive as in the other regions. The agro-ecological zoning invokes similarities in bio-geophysical conditions and land uses activities defining characteristics of surrounding ecosystem such as wetlands (Delepierre, 1982; Ndayambaje, Heijman, & Mohren, 2012). The exclusion of climate variables can be explained by the fact that the structural relief has an impact on rainfall distribution and hydrological network. On the Kivu lakeside, the Cyangugu backside and the Congo Nile Watershed Divide, slope elevation and contributing area are sufficient to predict wetland. These three agro-ecological zones are the result of geologically recent tectonic movement (Grzybowski, 1992) creating the Albertine Rift and the mountains with sharp boundaries. Here wetlands are bottomland receiving enough water from the well-watered upland. The region is the water tower of the whole country. Secondly, the eastern savanna, have relatively large swamps communicating with a series of lakes. The latter receives water from Akagera river overflow during the wet season, in a hydrological connectivity mechanism as described by Karim et al. (2012) for a similar case in Australia.

Our analysis shows that the central part of Rwanda, hilly with medium elevation and many small wetlands, has a potentially most climate change-sensitive wetlands. That is in line with others studies recognizing the changing impact of drivers with a variety of spatial scale of given ecosystem (de Koning, Veldkamp, & Fresco, 1998; Kok & Veldkamp, 2001).

Our study also illustrates that future-proof land use planning requires a multi-scale approach. National level land use plans need to be refined for specific targeting of climate sensitive wetland areas are strongly recommended.

Wetlands are sensitive to climate variations

The presented climate change scenarios demonstrate that wetlands in particular zones in Rwanda are more sensitive to climate change. This result

confirms the issue of lakes/wetlands sensitively expanding, reducing or even disappearing in some part of the country with climate variations (Nzigidahera, 2007). Our study also demonstrates that not all wetlands are equally sensitive to climate change and that the more sensitive regions are found in central Rwanda.

Model performance and accuracy

Model performance is generally better at the more detailed sub-national (AEZ and watershed) scale level. The models show differences between the descriptive and predictive power of specific variables (Apan & Peterson, 1998; Galletti, Ridder, Falconer, & Fall, 2013). The overall accuracy reaching 86.22 % - 93% with field based dataset are a clear improvement compared to studies using topographic derivatives only. Hogg and Todd (2007), reached only 84% of accuracy in characterizing wetlands in southern Ontario using a very high-resolution DEM (2mx2m pixel size) at detailed scales.

The advantage of using empirical models based on statistical analysis instead of process based models is the case study sensitivity. Much of the local specificity is automatically incorporated by the use of local data. The main drawback is that many known system sensitivities based on known processes are not always incorporated or apparent. Process-based modeling has this capability, but usually, requires very specific data (mainly discharge and evaporation data) which is not commonly available in countries like Rwanda. We, therefore, consider our empirical approach a first step towards understanding the wetland probability distribution directing future process-based research.

Other ways to improve functional wetland characterization may be to use higher-resolution data (now not available) and/or other data such as accurate soil information. Given the human impact on wetlands (use for agriculture and the constructions of dams of drains) the inclusion of such factors could also improve wetland occurrence characteristic.

2.5 Conclusion

The results of this study demonstrate the relevance of combining terrain derivatives and climatic factors in characterizing functional wetland occurrence in Rwanda. In line with previous studies, the Rwandese case study confirms the explanatory power of topographic attributes (elevation, contributing area, and slope). Moreover, the logistic regression model demonstrates that climatic factors are also important in determining wetland occurrence probability. The importance of climatic factors varies across space and is more scale sensitive. The logistic regressions were able to quantify the significance of climatic conditions for wetland definition in Rwanda.

The multi-scale quantification of climate sensitivity, using different values for temperature and rainfall in the model illustrate that wetlands in Rwanda are highly sensitive to climate changes. An increase of 1% in the national average temperature can result in a reduction of more than 12.5% of current national wetland coverage. The multi-scale approach demonstrates that the wetlands in the central part of the country are the most sensitive to climate change.

Our results underline that future wetland planning should be based upon at a multi-scale approach to ensure future-proof utilization of this important natural resource in Rwanda.

Chapter 3: Schistosomiasis mansoni incidence data in Rwanda can improve prevalence assessments, by providing high-resolution hotspot and risk factors identification[2]

[2] This chapter is based on: Nyandwi, E., Veldkamp, T., Amer, S., Karema, C. and Umulisa, I. Schistosomiasis mansoni incidence data in Rwanda can improve prevalence assessments, by providing high-resolution hotspots and risk factors identification, BMC Public Health (***In review after second revision***).

Abstract

Schistosomiasis mansoni constitutes a significant public health problem in Rwanda. The nationwide prevalence mapping conducted in 2007-2008 revealed that prevalence per district ranges from 0 to 69.5% among school children. In response, mass drug administration campaigns were initiated. However, a few years later some additional small-scale studies revealed the existence of areas of high transmission in districts formerly classified as low endemic suggesting the need for a more accurate methodology for identification of hotspots. This study investigated if confirmed cases of schistosomiasis recorded at health facility level can be used to, next to existing prevalence data, detect geographically more accurate hotspots of the disease and its associated risk factors. A GIS-based spatial and statistical analysis was carried out. Confirmed cases, recorded at 376 health facilities, were combined with demographic data to calculate incidence rates for each health facility service area. Empirical Bayesian smoothing was used to deal with rate instability. Incidence rates were compared with prevalence data to identify their level of agreement. Spatial autocorrelation of the incidence rates was analysed using Moran's Index, to check if spatial clustering occurs. Finally, the spatial relationship between schistosomiasis distribution and potential risk factors was assessed using multiple regression. Incidence rates for 2007-2008 were highly correlated with prevalence values (R2= 0.79), indicating that in the case of Rwanda incidence data can be used as a proxy for prevalence data. We observed a focal distribution of schistosomiasis with a significant spatial autocorrelation (Moran's I > 0: 0,05 – 0.20 and p ≤ 0,05), indicating the occurrence of hotspots. Regarding risk factors, it was identified that the spatial pattern of schistosomiasis is significantly associated with wetland conditions and rice cultivation. In Rwanda the high density of health facilities and the standardized microscopic laboratory diagnostic allow the derived data to be used to complement prevalence studies to identify hotspots of schistosomiasis and its associated risk factors. This type of information, in turn, can support disease control interventions and monitoring.

Keywords: *Schistosomiasis mansoni*, incidence rates, risk factors, spatial scale, empirical model.

3.1 Introduction

Schistosomiasis remains one of the most prevalent water-based diseases in the tropics. Regarding the impact, it is considered the second most important parasitic disease after malaria in many countries in sub-Saharan Africa (Sacko et al., 2011; Schur et al., 2013). In Rwanda, *schistosomiasis mansoni*, (written as "*S. mansoni*" in this chapter) with district level prevalence ranging from 0 to as much as 69.5% among school children constitutes a significant public health problem. The overall country prevalence in 2007-2008 was 2.7%.

The Neglected Tropical Disease (NTD) control program was established in 2007 by the Ministry of Health to fight against five NTDs which pose a significant public health problem. This program included the Schistosomiasis Control Initiative, which was implemented using the nationwide school-based prevalence map of 2007-2008 as a guideline. The prevalence recorded at 2 to 4 surveyed schools per district was averaged to estimate prevalence at the district level (nationwide 136 schools were surveyed). The Mass Drug Administration (MDA) for *S. mansoni* targeted children in areas with a prevalence of at least 10% and included adults where prevalence exceeded 30% (WHO, 2006).

However, two years later, some health facility located in districts classified as having a low prevalence, recorded higher frequencies of *S. mansoni* infection. In addition, some small-scale prevalence surveys revealed the existence of localized geographic areas with higher prevalence (up to 77.9%) than previously reported (Ruberanziza et al., 2015; Ruberanziza et al., 2010). One such study identified a 'new' area with the high rates of schistosomiasis (Isabwe et al., 2012), which was not detected before by the national prevalence study. The latter confirms the general recommendation of prevalence based studies in Rwanda or elsewhere in Africa (Meurs et al., 2013; Standley et al., 2009) that more detailed information is required to address the often highly focalized spatial pattern of schistosomiasis hotspots. The same studies also recommend to include other high-risk community groups (e.g.: women of children bearing age, rice farmers, fisherman) in future investigations.

To achieve this, the nationwide school-based prevalence surveys would need to include a very large number of schools and other high-risk community groups to provide sufficient information to identify and delineate hotspots. In Rwanda, districts are relatively large administrative units (see Figure 11a). To overcome the low granularity of the current mapping method (Congdon, 2009), the prevalence surveys would require large numbers of sample locations and becomes very expensive and harder to execute. Consequently,

there is a clear need to explore the value of other sources of health data and alternative mapping approaches to complement prevalence inventories and support planning and implementation of *S. mansoni* control programs. For other water based diseases, incidence data from routine health statistics have been successfully used to complement prevalence studies and identify the spatial distribution of a given disease (Odoi et al., 2003; Osei & Duker, 2008; Vega-Corredor & Opadeyi, 2014). Given the availability of good quality spatially structured and systematically collected data at the health facility level, use of such incidence data seems feasible in Rwanda.

Since incidence and prevalence have a direct relationship (prevalence = incidence rate x average duration of disease), a logical first step is to explore to what extent both have a similar spatial distribution. If there is a significant level of the agreement, it can be useful to use the much higher resolution incidence data for improved mapping of hotspots, and so more efficiently guide control interventions to high-risk locations.

From previous studies, it is known that the transmission of *S. mansoni* follows complex pathways depending strongly on the dynamics of the local biophysical and socio-economic context. Socio-economic and biophysical factors together influence when and where humans are in contact with water potentially contaminated with freshwater snails (Hu et al., 2013). In the Rwandan situation, new infrastructural developments such as dams and irrigation scheme expansion could very well contribute to the spread of *S. mansoni* to previously non-endemic areas (Utzinger et al., 2009). The association with irrigation projects is well scrutinized by Steinmann, Keiser, Bos, Tanner, and Utzinger (2006) in their systematic review of the relation between schistosomiasis occurrence and irrigated areas in some African countries. Specific socio-economic conditions such as educational attainment, access to improved water sources, and proper sanitation, are also known to be related to *S. mansoni* infection in endemic areas (Rollemberg et al., 2015; Schmidlin et al., 2013).

This study aims to investigate if schistosomiasis incidence data recorded at health facility level can provide, next to existing prevalence data, additional insights into the spatial pattern of schistosomiasis occurrence. Given the fine-grained spatial resolution of health facility service areas in Rwanda, this could allow more detailed spatial hotspot detection of the disease and its associated risk factors.

3.2 Materials and methods

Study area

Rwanda is a relatively small landlocked country of 26,338 km^2 in the Great Lakes region of Central-Eastern Africa with a climate characterized by two rainy and two dry seasons. Administratively, Rwanda is divided into five provinces, 30 districts and 416 sectors (MINALOC, 2005). The sectors together make up 376 health facility service areas (HFSAs), as shown in Figure 11b. Around 11 million people inhabit the country (NISR, 2014), of which 83.5 % live in the countryside mainly engaged in small-scale farming. Because of shortages of agricultural land, wetland conversion is one of the ongoing activities for rural development (Nabahungu, 2012).

Over time, Rwanda's health sector has experienced a profound evolution from traditional healing methods to faith-based health care during the colonial period, then centralized, and free provision of health services up to the early 1990's, to the current decentralized health care delivery system. The decentralization reinforces community participation in the management and financing of health services. The current public health care delivery system is a hierarchically organized three-tier system providing primary, secondary and tertiary health care.

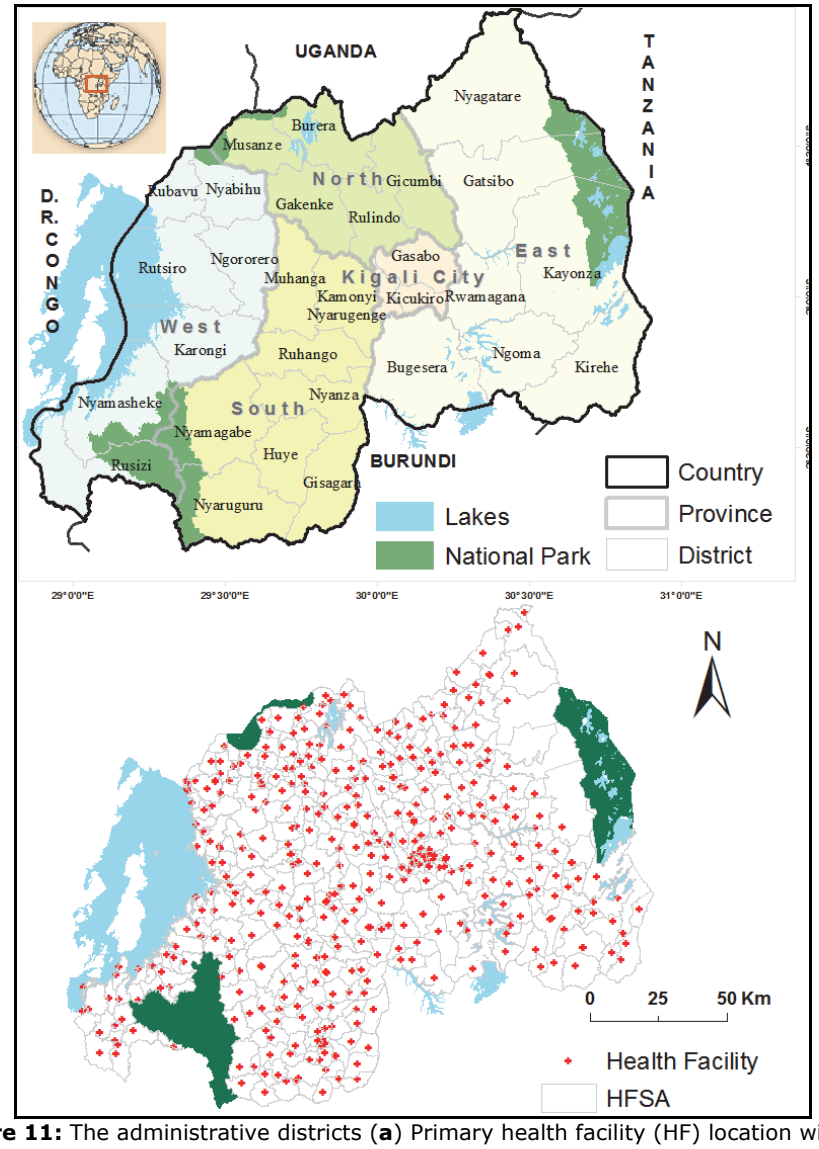

Figure 11: The administrative districts (**a**) Primary health facility (HF) location with their respective service area (HFSA) boundary (**b**).

Primary health care is provided at sector level via one or more health facilities, secondary care is provided in the districts by a district hospital, and referral hospitals provide tertiary care at the regional scale. Each health facility reports to the district health office, which is responsible for the health facilities and services provided to the population of the district. Community Health Workers (CHWs)function as an effective link between communities and public health care initiatives (Health Policy of Ministry of Health elaborated in 2005).

Data collection and quality checking

Administrative boundaries and spatial delineation of health facility service areas

The geographic location of health facilities depicting the 2007 situation, were provided by the Rwanda Biomedical Centre - Malaria and Other Parasitic Diseases Division (RBC/M&OPD). For this study, the health facility data were updated in May-June 2013 via intensive consultation of District Land Officers, Surveyors, and GIS Technicians from each administrative district. Furthermore, we spatially demarcated the HFSA using the 'cost allocation' spatial analyst tool of ArcGIS. Then, the delineated areas were further adjusted to the administrative units considering the population size, physical and managerial boundary as planned for Community-Based Health Insurances (CBHI) scheme management (In Ministerial Instruction Nr /Min/2012 on District - Health -Guidelines, 2012). The administrative boundaries of the country at different levels were obtained from the National Institute of Statistics of Rwanda. The general information data sets such as lakes, islands, parks, and roads, have been acquired from the Centre for GIS and Remote Sensing of the University of Rwanda (CGIS-UR).

Schistosomiasis mansoni incidence data

The number of confirmed cases of *S. mansoni* for the years 2007 - 2012 were provided for this study by the RBC/M&OPD. The same department also provided point locations and prevalence data obtained at the schools included in the nationwide school-based prevalence survey of 2007-2008. Another prevalence map produced using supra-national data, and regional simulation was also available from the World Health Organization (Chitsulo et al., 2000). The quality of recorded schistosomiasis infection cases at HFSA level is sound for six reasons. First, *S. mansoni* infection is diagnosed via Microscopic identification of eggs in stool samples in the laboratory of the health facility. Second, accessibility is unproblematic as patients can reach a health facility within a walking distance of at most 5 kilometres. Third, Community-Based Health Insurance (CBHI) makes that appropriate health care is affordable for everyone (patients pay only 10 % of the total cost of service and medication). Fourth, community health workers actively stimulate patients to visit the health facility in case of suspected health problems. Fifth, since there is no traditional medicine used for Schistosomiasis in Rwanda, patients with symptoms will go for treatment at the health facility. Sixth, the Rwandan Health Management Information System (R-HMIS), routinely and systematically collects health records from individual health facilities using a web-based software platform (DHIS2) whereby each health facility enters their monthly health records directly into the national database (USAID/Rwanda, 2006).

Socio–economic and biophysical covariates

Demographic data were extracted from the 2002 and 2012 Population, and Housing Census published by the National Institute of Statistics of Rwanda (NISR) and for the years in between the estimated population number and distribution were extracted from estimation and projections also reported by NISR. The socio-economic factors such as school attendance levels, access to improved water sources, proper sanitation at district and HFSA level were available from census reports.

S. mansoni transmission is determined and accelerated by interactions of various factors spatially restricted to freshwater bodies inhabited by particular host snails (Walz, Wegmann, Dech, Raso, & Utzinger, 2015). Earlier studies in East African countries identified numerous biophysical and socio-economic conditions related to schistosomiasis infection risk (Schur et al., 2013). In addition to socio-economic factors, biophysical factors need to be considered as well. Wetland agro-ecosystems related factors (namely wetland proportion, rice cropped areas, wetland/water body adjacency) were also collected. Likewise, topographic and climatic factors were used as potential risk factors for this study. Data acquisition and pre-processing methods to generate raster data for the risk factors are detailed in previous research on Rwandan wetlands characterization and their climate sensitivity (Nyandwi, Veldkamp, & Amer, 2015). Soil parameters (pH, clay, and sand content percentage) were extracted from the soil geo-database of Rwanda generated from a semi-detailed soil survey consisting of 1833 soil profiles spread over the country. The geostatistical interpolation of soil properties was done using landform data at a scale of 1: 250 000 and 1: 50 000 (Verdoodt & van Ranst, 2006). Mean values at HFSA level were extracted from original risk factors raster data using zonal statistics tools of ArcGIS 10.2.2. District Factor values are the average of values of HFSAs within a district.

Comparison of spatial distribution of *S. mansoni* with incidence and prevalence datasets

Incidence and prevalence are both measurements of disease frequency. Incidence estimates how often disease occurs in space and time (a measure of disease risk). Prevalence evaluates how much the disease is spread in a given population (a measure of disease burden) at a given moment. Since both are related (prevalence = incidence rates x average duration of disease), high prevalence areas may correspond with high incidence rates for a disease such as *S. mansoni.* The 2008 prevalence at each of the 136 surveyed schools is directly compared with incidence rates at HFSA level for the same year and location. Prevalence and incidence data were also mapped at the district level to enable visual comparison.

Detection and visualization of the spatial pattern of S. mansoni

Appropriate spatial and statistical approaches for detecting spatial clustering and relationships with risk factors are provided by current advances in spatial epidemiology (Osei, 2014; Tsai, Lin, Chu, & Perng, 2009).

Cases of S. mansoni per population at risk

The number of confirmed cases of *S. mansoni* per HFSA per month were summed per year and further aggregated to the district level. The generated annual *S. mansoni* incidence cases and demographic data were joined using Excel and then linked to the HFSA and district spatial data. The incidence rate is the number of *S. mansoni* cases per district (n = 30) or HFSA (n = 367) divided by the population of that district or HFSA, as shown by the equation below:

$$Di = (In / Pt) * 100\,000 \qquad \textbf{\textit{(Eq. 3)}}$$

Where ***Di*** is the S. mansoni incidence rate, ***In*** is the total number of new cases in 12 months of a year per district/HCSA, and ***Pt*** is the total population of that year for that entity.

To make the rates more intuitive, they were multiplied by 100 000 to obtain incidence rates reported per one hundred thousand persons (Rothman, 2012). The raw rates from sparsely populated HFSAs were replaced by weighted averages using Empirical Bayesian Smoothing (EBS) (Goovaerts, 2008). EBS computes raw rates and produces three weighted rates using the global, mean and local average. The smoothing was done using the SpaceStat software (CDC, 2014).

For visualization, incidence rates were classified into four classes using the Jenks classification. The classes defined in this way were harmonized to allow for inter-annual comparison and comparison with prevalence maps at the district level. Furthermore, the average incidence rates for 2007 and 2008 at HFSA level were standardized to vary from 0 to 1 (from non-endemic to hyperendemic areas) and superimposed with the points map of the 136 schools surveyed during the nationwide mapping of 2007-2008. A scatterplot was then generated to display the relationship between incidence rates at HFSA level and prevalence per school.

Analyzing pattern with Moran's Index statistic

Spatial autocorrelation was computed to ascertain the correlation between neighboring incidence rates of *S. mansoni* and the level of spatial clustering within the study area (Boots & Getis, 1998). The Moran's Index statistic, similar to the Pearson correlation (Cliff & Ord, 1973), is widely used for this and was calculated as:

$$I = \frac{N}{S_O} \sum_i \sum_j w_{ij} \frac{(x_i - u)(x_j - u)}{\sum_i (x_i - u)^2},$$

(Eq.4)

Where ***N*** is the number of districts/HFSA; $\boldsymbol{W_{ij}}$ is the element in the spatial weights matrix corresponding to the observation pair ***i, j.*** Also, $\boldsymbol{x_i}$ and $\boldsymbol{j_i}$ are observation for areas ***i*** and ***j*** with mean ***u***. And

$$S_O = \sum_i \sum_j w_{ij}$$

(Eq.5)

Since the weights are row-standardized $\sum \boldsymbol{w_{ij}} = 1$, the first step in the spatial autocorrelation analysis is to construct a spatial weight matrix that contains information about the neighbourhood structure for each location. Adjacency is defined as the immediately neighbouring district/HFSA, including the district/HFSA itself (Anselin, 1995). Non-neighbouring units have a weight of zero.

Mapping clusters

The Local $G_i^*(d)$ statistic was selected to test the statistical significance of local clusters and to determine the spatial extent of these clusters (Getis & Ord, 1992; Ord & Getis, 1995). The Local $G_i^*(d)$ statistic is useful for identifying individual members of local clusters by determining the spatial dependence and neighbouring observations (Tsai et al., 2009; Wu et al., 2004). It can be written as follows:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d) x_j - W_i \bar{x}}{s \sqrt{\frac{\left(n S_{1i} - W_i^2\right)}{(n-1)}}}, \text{ for all } j$$

(Eq.6)

Where x is a measure of incidence rate of S. mansoni within a given district/HFSA polygon; W_{ij} is a spatial weight that defines neighbouring district/HFSA j to i; W_i is the sum of the weight W_{ij},

$$\bar{x} = \frac{1}{n} \sum_j x_j \quad S_{1i} = \sum_j w_{ij}^2, \quad s^2 = \frac{1}{n} \sum_j x_j^2 - \bar{x}^2 .$$

(Eq.7)

Developing the spatial weight W_{ij} is the first step to calculating Gi^* (d). The spatial weight matrix includes W_{ij} = 1. and in this study, the adjacency has been defined in ArcGIS - proximity analysis based on polygons that share common boundaries and vertices (Ceccato & Persson, 2002).

With Local $G_i^*(d)$ statistic clusters with a 95 percent significance level from a two-tailed normal distribution indicate significant spatial clustering, but only positively significant clusters are mapped.

Empirical modeling of *S. mansoni* with potentially associated factors

Input data sets were standardized and prepared at HFSA level using standard functionality of ArcGIS. The attributes of the spatial data were exported to IBM SPSS Statistics, version 20 which was used for all statistical analyses. Before the use of the variables to train an exploratory empirical model, exploratory data analysis was done. Considering the patterns of data distribution, using normal distribution curves and frequency distributions; a log-normal distribution transformation was done for some of the input datasets (Limpert, stahel, & Abbt, 2001). Then, with normal and ln-transformed variables, Pearson's rank correlation coefficient and the test of co-linearity using pairwise scatter plots was done. Sixty percent of the data was randomly sampled for the model, to avoid the spatial autocorrelation in our empirical regression and to prevent a fixed spatial pattern. With the spatial autocorrelation, the model becomes insensitive to changes in the spatial patterns and over fitted.

The incidence of *S. mansoni*, as the dependent variable, was related to all potential risk factors. Risk factors used in the statistical analysis were grouped into five. The first group is made up of physical variables, sub-categorized as soil properties (pH, sand percentage, and clay percentage) and terrain derivatives (elevation, slope, and terrain shape index). The second group represents ecological variables (wetland proportion, the total area of rice cropping schemes, water/lake adjacency). The third group consists of two climatic variables (total annual amount of rainfall and average annual temperature). The fourth group consists of four demographic variables (number of households, population density, and percentage of rural and urban residents). The last group represents socio-hygienic conditions, namely: level of education, the source of water and sanitation situation.

A stepwise linear regression was conducted, considering $P > 0.1$ as the removal criterion and $P \leq 0.05$ as the entry criterion. This quantifies the strength of the relationship between schistosomiasis cases and the significant covariate(s). The standardized coefficients were used to compare the effects of each independent variable on the dependent variable (Lemeshow & Moeschberger, 2005). To determine how well the regression model fitted the data, we used the R-square and the Standard Errors of the Regression (S). We considered the S to represent the average distance that the observed values deviate from the regression line. The S must be ≤ 2.5 to produce a sufficiently narrow 95% prediction interval (Kabo, Paulechka, Zaitsau, & Firaha, 2015). Also, we used both residuals and residual plots to analyze drift

and variance of the values (Wang et al., 2015), for evaluating the appropriateness of the model.

3.3 Results

Comparing Prevalence and Incidence data

To check if indeed there is a significant relationship between prevalence and incidence data, we compared standardized incidence rates of 2007-2008 at HFSA level with prevalence data obtained at the 136 schools sampled during the 2007-2008 nationwide survey. In Figure 12, we can observe a very strong relationship between the two data sources with a coefficient of determination of 0.79.

This strong relationship indicates that we can use routinely recorded cases of *S. mansoni* at primary health facility level to supplement prevalence data. The readily available incidence rates can be utilized as a proxy for prevalence. This information can be sourced on a monthly basis for every health facility for more than one decade.

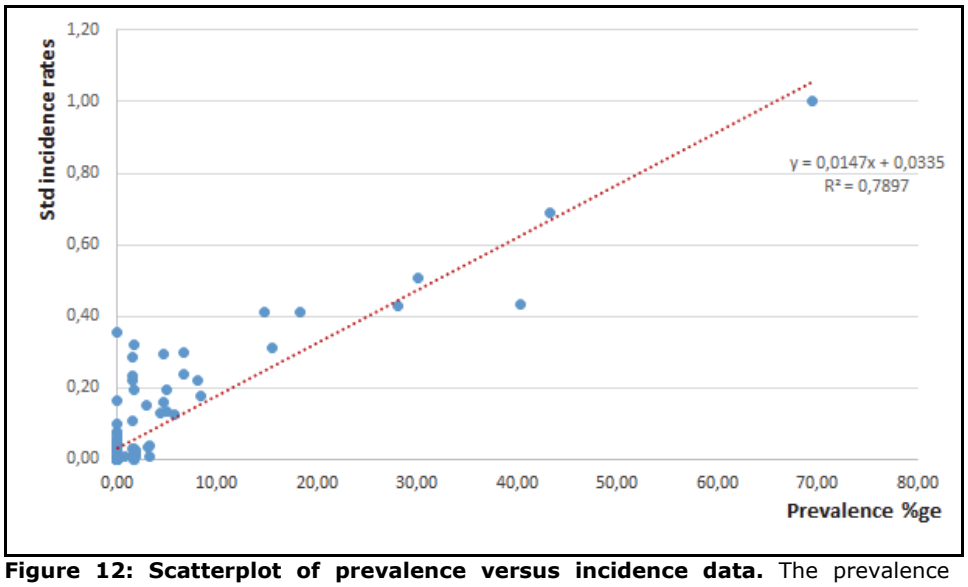

Figure 12: Scatterplot of prevalence versus incidence data. The prevalence proportion measured at each of 136 surveyed schools linked to their location; and corresponding standardized incidence rates were extracted for the same places.

Spatial distribution of S. mansoni in Rwanda

A total number of 1221 *S. mansoni* cases were reported in Rwanda for the years of 2007 and 2008. Annual incidence rates for 30 districts were calculated (Table 6). Nyagatare district had the highest rate of 57 cases per 100 000 persons in 2008. Then EBS incidence rates for 367 HFSAs were also calculated for the year 2007 and 2008. As visualized in Figure 13, a lot of

HFSAs, as well as some districts, have zero or very low numbers of *S. mansoni* cases per year. Visually, there also are districts with high rates (respectively Nyagatare, Kirehe, Ngoma, Rusizi, and Burera); districts with zero rates (10 Districts in 2007 and only 2 in 2008) and districts with very low incidence rates. At HFSA level, the spatial pattern of *S. mansoni* is much more distinct, showing considerable differences within a district.

Table 6. Incidence rates (per 100.000 persons) of *S. mansoni* in Rwanda for specific years (2007, 2008 and cumulative incidence (2007-2008)

District Name	**Year**		
	2007	**2008**	**2007-2008**
Nyarugenge	2,33	2,67	3,82
Gasabo	5,39	6,42	8,40
Kicukiro	0,41	1,17	0,78
Nyanza	0,38	7,05	3,71
Gisagara	1,05	1,37	1,72
Nyaruguru	0,39	1,51	1,34
Huye	0	11,43	5,71
Nyamagabe	0	0,64	0,32
Ruhango	0,72	0	0,71
Muhanga	0,66	0,98	1,31
Kamonyi	0	0	0
Karongi	1,66	0,33	1,96
Rutsiro	8,64	7,13	12,22
Rubavu	0	0,58	0,29
Nyabihu	3,92	6,7	7,05
Ngororero	0	4,2	1,94
Rusizi	17,91	21,66	28,42
Nyamasheke	4,57	15,46	12,37
Rulindo	0,75	1,47	1,47
Gakenke	0	1,81	0,91
Musanze	3	7,96	7,08
Burera	16,45	47,6	40,02
Gicumbi	2,65	8,4	6,83
Rwamagana	0	18,83	9,60
Nyagatare	1,57	**57,07**	**30,04**
Gatsibo	0	10,35	5,18
Kayonza	0	1,13	0,76
Kirehe	**35,25**	16,89	42,40
Ngoma	10,98	2,13	11,72
Bugesera	0	1,28	0,64

Spatial autocorrelation of *S. mansoni* rates

The results of spatial autocorrelation of neighbouring values aggregated at HFSA and district level are summarized in Table 7.

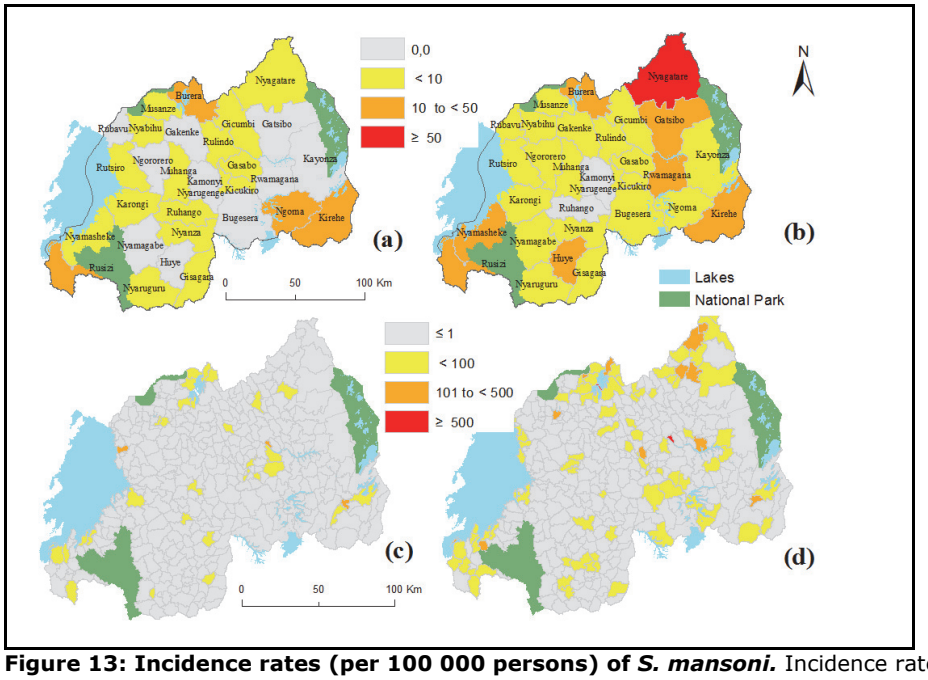

Figure 13: Incidence rates (per 100 000 persons) of *S. mansoni.* Incidence rates at the district level for 2007 (**a**) and 2008 (**b**) and EBS incidence rates at HFSA level for 2007 (**c**) and 2008 (**d**).

The results were statistically significant at the district level and strongly significant at HFSA level ($p < 0.05$ and z-score greater than 1.96). The statistically significant values indicate that the distribution *S. mansoni* is spatially heterogeneous in Rwanda and that heterogeneity is more explicit at HFSA level.

Measures of spatial clustering (hotspots) of *S. mansoni* rates

The identified statistically significant hotspot areas from the Local Gi* (d) test of *S. mansoni* rates for the year 2007 and 2008 are visualized in Figure 13.

Table 7: Test of spatial autocorrelation of *S. mansoni* rates computed for cumulative incidence (2007-2008) and the particular years (2007and 2008)

Year	District level		HFSA level	
	Moran'I (p value)	Z (I)	Moran'I (p value)	Z (I)
2007-2008	0.13 (<0.001)	5.03	0.14 (<0.001)	5.03
2007	0.24 (0.05)	1.89	0,12 (<0.001)	4.36
2008	0.12 (0.29)	1.10	0.11(<0.001)	3.65

The outcomes from spatial clustering analysis computed with Local *Gi* (d)* statistic at district and HFSA levels are categorized as clusters (z-score ≥ 1.96) or non-clusters (z-scores < 1.96), at different significance level.

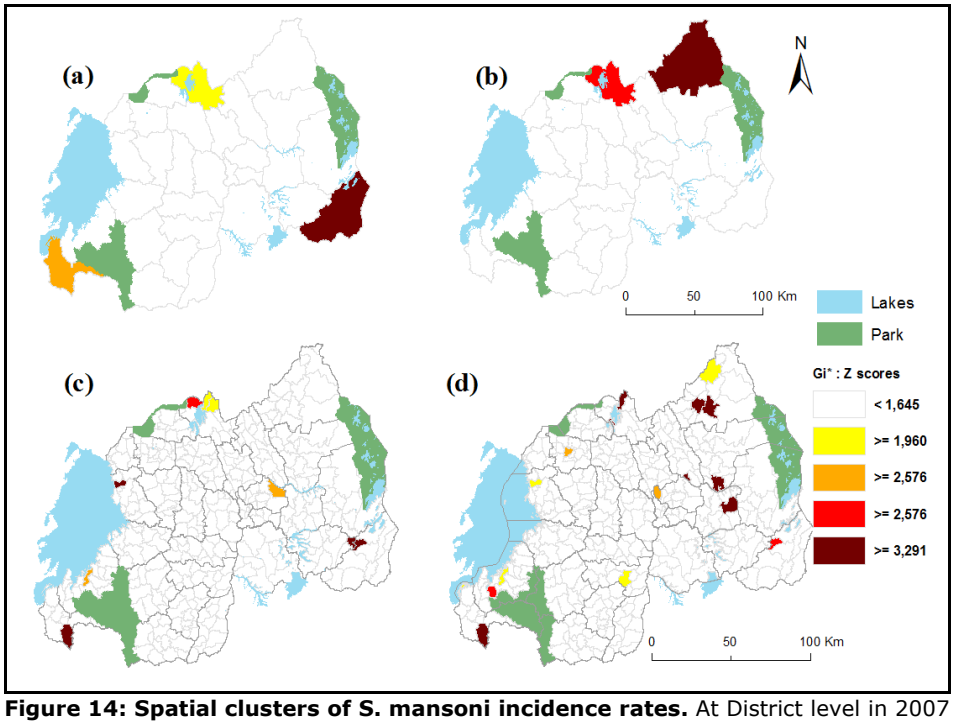

Figure 14: Spatial clusters of S. mansoni incidence rates. At District level in 2007 (**a**), in 2008 (**b**) and at HFSA level in 2007 (**c**), in 2008 (**d**). These maps show five levels of statistical significance of Z-score values: Not important spot, has a value < 1.645. Hot Spot with 90% confidence ≥1.645; follows ≥ 1.960 for Hot Spot with 95 % Confidence ≥ 2.576 Hot Spot with 99% Confidence and ≥ 3.291 Hot Spot with 99,9% Confidence.

S. mansoni spatial distribution in relation to environmental factors

Some risk factors have a significant relationship with *S. mansoni* incidence. The sand percentage (of the soil) and elevation are negatively correlated with *S. mansoni* incidence rates, while ln-terrain shape index (Hurlimann et al.), temperature, rain, and rice cropped area in wetlands is positively correlated (see Table 8). Using R-squared and Standard error of the estimate, the risk factors included in the empirical model explain quite a lot regarding the spatial distribution of *S. mansoni* with a distinct effect of spatial scale. More than 47% of the distribution (with S of 0.926) at detailed HFSA level and 60 % (with S of 0.366) at the larger District level.

Table 8. Multiple regression outputs for the relationship between environmental factors and *S. mansoni* at HFSA and District level.

Parameters	B Coeff.	Standard Error	Beta Coeff.	t-value	p-value
HFSA level model					
Intercept	-0.697	0.238	-	-2.931	0.004
Sand percentage	-0.004	0.001	-0.366	-3.816	0.000
Rice cropped area	0.001	0.000	0.256	2.692	0.009
log-TSI	0.131	0.041	0.298	3.187	0.002
Rain	0.001	0.000	0.281	2.585	0.012
Temperature	0.028	0.010	0.296	2.832	0.006
District level model					
Intercept	1.999	0.906	-	2.207	0.036
Sand percentage	-0.032	0.010	-0.447	-3.216	0.003
Elevation	-0.002	0.000	-0.661	-4.373	0.000
Rain	0.002	0.000	0.657	4.336	0.000

None of the two models included socio-hygienic variables such as educational attainment, the source of water or sanitation conditions. Even with the univariate test by Pearson's rank correlation (not reported here), none of the hygienic and socio-demographic factors had a significant association with *S. mansoni* incidence.

3.4 Discussion

This study identifies and visualizes the spatial variability of *S. mansoni* at two levels of spatial resolution using routinely collected health records as a basis. The incidence rates generated at HFSA levels were EBS smoothed, and the global mean was able to assign new rates, as recommended for disease mapping at a high spatial resolution (Bithell, 2000) and in line with the considerable geographic concentration of *S. mansoni* in Rwanda (Ruberanziza et al., 2015). The correlation between neighboring values at small scale was supported by the Global Moran's index. The spatial clustering test using Local $G_i^*(d)$ statistic also shows the non-random spatial distribution of *S. mansoni*.

Routine health records provide valuable information for spatial pattern detection of S. mansoni

Disease maps depicting the spatial pattern of *S. mansoni* are essential for guiding control program activities. However, the added value of disease maps much depends on their spatial resolution, and on the underlying data used to establish them (Moore & Carpenter, 1999). In this section, we first compare the incidence rates based maps at district and HFSA levels. The second

comparison is between incidence rates based mapping and prevalence based mapping.

Figure 15 illustrates how spatial scale influences the detection of disease hotspots. The centre map of Figure 15 depicts incidence rates at the district level for 2007-2008, while the four smaller maps show the same information but now at the HFSA level. The HFSA level maps clearly show the highly focalized nature of *S. mansoni* hotspots. Representation at the district level, on the other hand, results in an overestimation of areas of high transmission as well as in non-identification of hotspots in districts with generally low incidence rates. The four identified hotspots areas at HFSA level (see Figure 5) have also been identified by previous studies. The first and second hotspot areas are historically endemic zones of *S. mansoni*. Recently, Ruberanziza et al. (2015) reported Nkombo Island (hotspot **1)**, as the most important *S. mansoni* focus in Rwanda. Ntaruka HFSA (hotspot **2)** between Burera and Ruhondo lakes, was previously also identified as a high transmission area by several cross-sectional surveys (Hanotier & Gigase, 1981; Mupfasoni et al., 2009; Ruberanziza et al., 2010). Hotspot **3** in Nyagatare and hotspot **4** in Gasabo district have not been documented before.

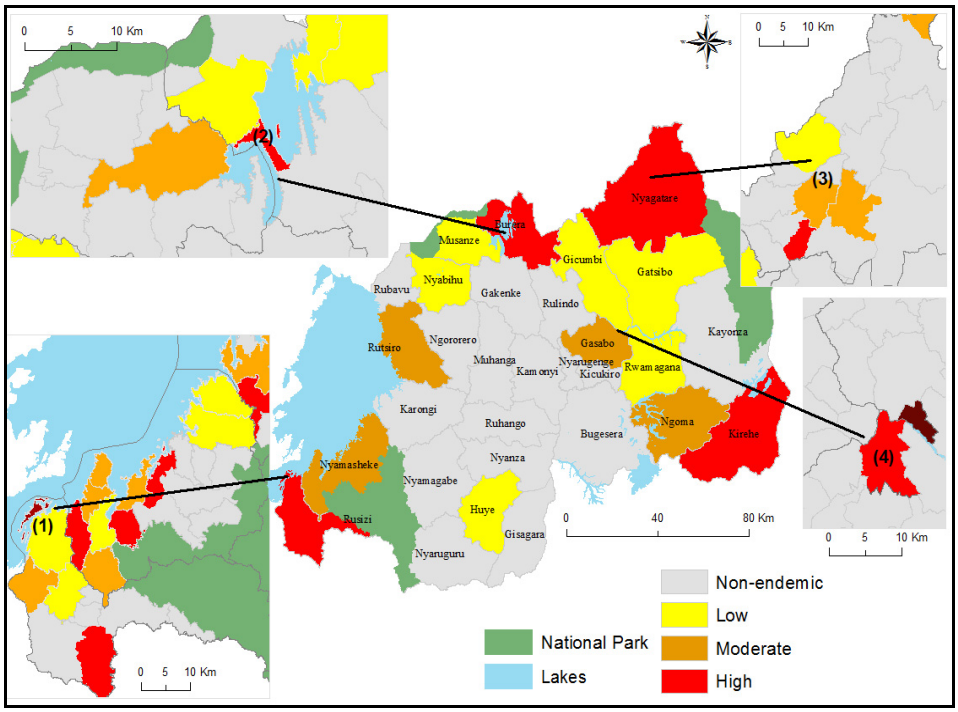

Figure 15: Spatial scale sensitivity of *S. mansoni* incidence rates. The large map in the center shows incidence rates (2007-2008) per district. This is then compared with rates displayed at HFSA level using known hotspots areas as a reference.

The second comparison, illustrated in Figure 16, is at district level between the incidence-based map and two prevalence-based maps. The first (see Figure 16a) is the prevalence map of 2008 published by WHO in 2010 (2010a). The WHO map was produced using United Nations population data and prevalence estimations based on the procedure developed by Chitsulo et al. (2000). If we compare the WHO map with the one based on incidence data (Figure 16b) there are obvious similarities, but with a notable exception for the Gicumbi district (red unit in Figure 16a). According to the WHO mapping, Gicumbi district is hyper-endemic while in reality, the nationwide school-based prevalence survey (see Figure 16c) identified three out of four sampled schools to be non-endemic (with a prevalence of 0%).

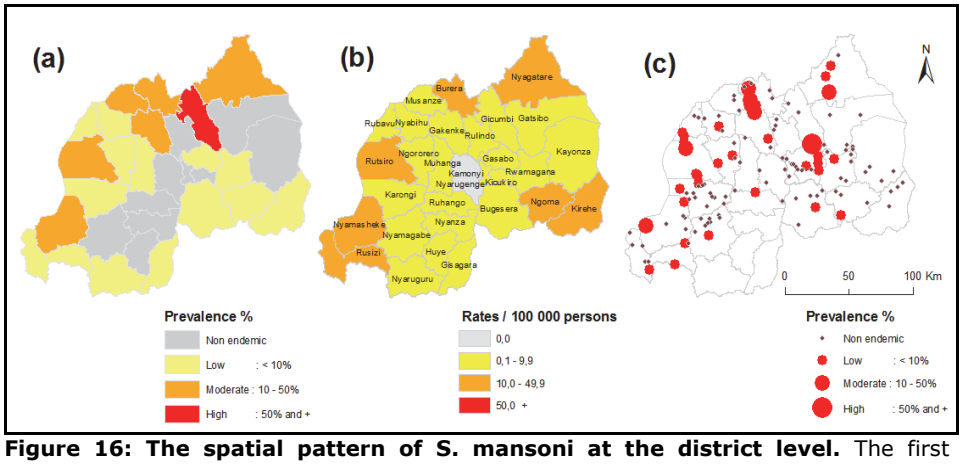

Figure 16: The spatial pattern of S. mansoni at the district level. The first prevalence map of 2008 was produced by WHO using global data and spatial simulation (**a**); the centre map shows the cumulative incidence rates (2007-2008) at the district level to allow comparison (**b**) and the school-based prevalence survey of 2007- 2008 (**c**).

In order to obtain a reliable disease burden measure, a prevalence survey is always needed. But the use of incidence data is suitable to identify high-resolution spatial patterns of disease distribution at the national scale. Disease maps based upon incidence data might also be used to guide a spatially explicit sampling procedure for improved prevalence sampling. This is important given that the spatial representability of surveyed schools is not well elaborated in current WHO guidelines for the evaluation of helminthiasis at the community level (Montresor, Crompton, Hall, Bundy, & Savioli, 1998). In some case, the cross-sectional survey was influenced by accessibility by four-wheel drive car (Assare et al., 2015). Inaccessible areas with poor quality or non-existent roads around valleys and perennial water bodies are usually poorly represented in school samples obtained, while those have now been identified as potential high-risk areas (Simoonga et al., 2008).

***S. mansoni* incidence rates and environmental risk factors**

This study detected a significant relationship between *S. mansoni* incidence rates and potential environmental risk factors as summarized in Table 8. Elevation and sand percentage in the soil are negatively correlated with *S. mansoni* incidence, whereas temperature, rainfall, and wetlands used for rice cultivation are positively associated. These findings are consistent with existing knowledge on the environmental conditions required for the development of the intermediate host snails and hence, can influence transmission risk. The related physical and ecological parameters were similar to other studies (Alemayehu & Tomass, 2015; Fonseca, Freitas, Dutra, Guimarães, & Carvalho, 2014; Schur et al., 2013). Low elevation, valley wetlands, warm and humid conditions and less sandy soil create favorable conditions for *S. mansoni*, as host snails inhabit places with lower altitudes and wetlands with natural or cultivated vegetation (Alemayehu & Tomass, 2015; Bavia, Hale, Malone, Braud, & Shane, 1999). The TSI offer additional and important information for discriminating the implications of those changes on wetness conditions more than soil types do (Fontaine et al., 2007).

Prior research has demonstrated that *S. mansoni* infection can increase as a result of the construction of dams or irrigation schemes (Steinmann et al., 2006). Indeed, within irrigation systems transmission is focal and primarily due to localized contamination of habitats with human excreta containing Schistosoma eggs (Boelee, 2000). The newly identified hotspot in Nyagatare district (hotspot 3 in Figure 15) is close to a more than 1000 hectares irrigation scheme for rice cropping. Furthermore, intense cultivation such as flowers and sugarcane within Nyabarongo wetland can be linked with the hotspot identified in Kigali city (see hotspot 4 in Figure 15). A similar situation was also reported for Minas Gerais State in Brazil (Guimarães et al., 2008). This indicates that vulnerability to *S. mansoni* is not limited to rural populations, school children or women of childbearing age but extends to entire communities.

Limitations of this research approach

Although the incidence data from routine health records have a high quality as described in the methodology section, the number of confirmed cases will still only be a fraction of all infected persons and will mostly concern patients manifesting clinical symptoms. If the incidence database could be enriched with additional information (e.g. intensity of infection in tested stools, the total number of investigated subjects and their age and sex) the outcomes of the analysis would become even more robust.

Somewhat unexpected is that none of the demographic and socio-economic variables had a significant contribution to explaining *S. mansoni* incidence

variability at HFSA level. A plausible explanation for this is that in Rwanda there is low variability in levels of school attendance, access to improved water sources, proper sanitation, and wearing of shoes (Rollemberg et al., 2015; Schmidlin et al., 2013). In Rwanda, conditions have much improved as a result of significant policy achievements in the last two decades (MINECOFIN, 2015; United Nations, 2000). Consequently, additional information about people's behavior would be required to further improve our understanding of *S. mansoni* risk factors. Possible examples are the habit of taking off shoes while working in the field, defecation in bushes when working on the land (and thus far from the home toilet), which may be important to better understand *S. mansoni* exposure (Brooker et al., 2009; Congdon, 2009; Schur et al., 2013).

The multiple regression models explain a lot of the observed spatial variability of S*. mansoni* incidence rates as a function of possible locational risk factors. The district model performed better than the model at the more detailed HFSA level. This is consistent with the fact that aggregation of data causes linearization contributing to overestimation in linear regressions (Kok & Veldkamp, 2011).

3.5 Conclusion

This study has demonstrated that in Rwanda prevalence and incidence data for *S. mansoni* are highly correlated. Given the availability of reported cases for each health facility in Rwanda, a high resolution spatially explicit statistical investigation of *S. mansoni* hotspots is feasible. The identified risk areas provide an appropriate basis to guide *S. mansoni* control programs at a much more detailed spatial scale than was possible before. In addition, the most important physical, ecological and climatic risk factors for *S. mansoni* transmission in Rwanda were identified. It was also shown that intensive agricultural use and transformation of wetlands for rice cultivation contributes to the spreading of *S*. *mansoni* into previously non-endemic areas. In line with environmental health impact monitoring and evaluation for wetland based development projects, a specific policy is required to address and reduce potential disease risk associated with rural development efforts. Finally, use of routinely collected incidence data opens the door for spatiotemporal analysis of *S. mansoni* and environmental risk factors which will vary in space and time.

Chapter 4: Spatial Patterns of Soil-transmitted Helminth Infections and Associated Environmental Risk Factors in Rwanda 2007-2008[3].

[3] This chapter is based on: Nyandwi, E., Veldkamp, T., Amer, S., Ruberanziza, E., Rujeni, N. and Umulisa, I. Schistosomiasis mansoni incidence data in Rwanda can improve prevalence assessments, by providing high-resolution hotspots and risk factors identification, PLOS Neglected Tropical Disease Journal (***In review***).

Abstract

Soil-transmitted helminth (STH) infections are parasitic diseases with significant public health impact, mostly analysed and reported in combination and based on cross-sectional prevalence surveys. However, recent research demonstrates that infection levels and spatial patterns differ between these three STH types. Incidence data of STHs including roundworm (Ascaris lumbricoides), whipworm (Trichuris trichiura) and hookworms per primary health facility for 2007-2008 were linked to spatially delineated primary health centre service areas using ArcGIS 10.2. The prevalence data per District for individual and combined STH infections from the 2007/2008 nationwide survey in Rwanda were obtained together with other health data. A comparison of reported prevalence and incidence data aggregated at the District level for 2007-2008 indicates that significant positive correlations exist for roundworm (R^2= 0.63) and hookworm (R^2= 0.27). Weak positive correlations were observed for whipworm (R^2=0.02) and the three STHs combined (R^2= 0.10). Spatial autocorrelation of single and combined STH incidence rates was detected with Moran's Index test. The incidence of roundworm and whipworm were found to be focalized with a significant spatial autocorrelation (Moran's I > 0: 0.05 – 0.38 and $p \leq 0.03$.), with high to very high incidence rates in some focal areas. In contrast, hookworm incidence is ubiquitous, randomly distributed (Moran's I > 0: 0.006 and p= 0.74), and with very low incidence rates. Furthermore, an exploratory empirical regression analysis was done to identify relationships between helminth infection incidences and potential environmental and socio-economic risk factors. Findings show that the spatial distribution of STH incidence is significantly associated with soil properties (sand proportion and pH), rainfall, wetlands and their uses, population density and the proportion of rural residents. The detected spatial patterns are important to direct STH prevention and control programs.

4.1 Introduction

Soil-transmitted helminths (STH) infect humans through contact with parasite eggs (*Ascaris lumbricoides and Trichuris trichiura)* or larvae (hookworm) that thrive in moist and warm conditions in soils, water, and edible plants (Lai, Zhou, Utzinger, & Vounatsou, 2013). The burden of STH infections is well known in Rwanda. After malaria, STH represent the most important parasitic diseases in terms of public health and economic impact (TRAC+, 2008). The countrywide mapping of STH conducted in 2007/8 and results from a simulation done by WHO (2010a) demonstrated that 27 out 30 districts of Rwanda have an overall STH prevalence greater than 50%. The prevalence mapping of 2007/2008 used a cross-sectional school-based survey including around 4 schools per District, and results were extrapolated at the District level. The resulting high and most evenly distributed STH prevalence per district, guided the repeated massive drug administration (MDA) to children aged 1-16 years, post-partum mothers, and other intervention measures (Ruxin & Negin, 2012).

In our knowledge, there has been no follow-up nationwide prevalence mapping, but monitoring of the spatial pattern of STH infection is needed to assess the impact of interventions and plan future interventions measures. Therefore, an efficient and inexpensive approach is needed that can be used to monitor the spatial distribution of STH transmission over time.

The prevalence study conducted in 2007/2008 has three main limitations. Firstly, the district level mapping approach is too crude to capture the considerable geographic variability of biophysical and socio-economic conditions associated with STH transmission. Secondly, the existing STH prevalence inventory targets school-aged children only, ignoring other population groups equally exposed to the infection. Thirdly, differences in STHs pathogenicity and mode of transmission (A. Abera & Nibret, 2014) are ignored. It has been demonstrated that the spatial distribution of different types of STH infections is very heterogeneous in neighbouring countries of Rwanda (Pullan, Kabatereine, Quinnell, & Brooker, 2010) and in other endemic areas in tropical and sub-tropical conditions (Kaliappan et al., 2013). Also, Clements, Deville, Ndayishimiye, Brooker, and Fenwick (2010) revealed different spatial patterns for different STH types in the East African Great Lakes region, with ubiquitous hookworm infection, and highly focal prevalence patterns of *Ascaris lumbricoides* and *Trichuris trichiura.*

STH infections are associated with biophysical features that influence egg and larval survival (Halpenny, Paller, Koski, Valdés, & Scott, 2013). Also, socio-economic factors associated with (limited) sanitation, hygiene, and behavioral factors can facilitate parasitic eggs ingestion (Strunz et al., 2014). The latter

may differ from one to another STH as these have slightly different life cycles and transmission trajectory (Larsen & Roepstorff, 1999), as summarized in Table 9.

Table 9. Life cycle and transmission trajectory of Soil-transmitted Helminths

Type of Helminth	Living conditions and human transmission
Roundworm *(Ascaris lumbricoides)*	Rural area, in a hot and humid area where contamination of the soil is high from faeces. Once in the environment, infective third-stage develops within the eggs after around three weeks and can survive in the ground for up to 20 years. Transmission occurs through ingestion of infective eggs located in contaminated soil, water, and edible plants (Cho, 2009).
Whipworm *(Trichuris Trichiura)*	Eggs from faecal deposited in the warm and humid soil can mature in the soil after around 10-14 days. Transmission is through the faecal-oral route (Arizono, Yamada, Tegoshi, & Onishi, 2012)
Hookworm *(Ancylostoma spp)*	Larvae hatch in soil from eggs after 24 hours of being laid in stool; approximately 24 hours later the worms molt into infective filariform larvae that are capable of penetrating intact skin. Transmission to humans usually occurs through bare feet on contaminated soil (CDC, 2013).

Incidence data on STH that is routinely collected at primary health facilities exists in Rwanda. Instead of executing prevalence studies on a regular basis, these incidence data could be used for spatial distribution analysis. A first case study on schistosomiasis mansoni incidences in Rwanda (Chapter 3) demonstrated a very high correlation between prevalence and incidence data (R^2= 0.79). An added value of the existing incidence data is that it allows detecting detailed spatial patterns using the 367 available health facilities service areas (HFSA) every year.

This study aims to determine whether incidence data can be used to detect the spatial distribution of each helminth type as well as combined STH infection. In addition, it assesses how the spatial variation of STH transmission is related to environmental risk factors. This assessment is essential for understanding the spatial variation of STH incidences in Rwanda and to contribute to the efficient planning of future disease control programs.

4.2 Materials and methods

Study area description and detailed health facility catchment areas used are similar to the one on chapter 3.

Data collection and structuring

STH incidence and prevalence data. Confirmed cases of STH infection, including roundworm (*Ascaris lumbricoides*), whipworm (*Trichuris trichiura*) and hookworms (*Ancylostoma duodenale* or *Necator americanus* – the available data do not distinguish between the two), recorded at primary health facility level for 2007-2008 were obtained from the Rwanda Biomedical Centre - Malaria and Others Parasitic Diseases Division (RBC/M&OPD). Prevalence data per District for individual and combined STH infections from the 2007/2008 nationwide survey were also obtained from this division. The period of 2007-2008 was selected for two reasons: *(i)* it corresponds with the period of nationwide school-based survey and mapping of STH prevalence (2007/2008), *(ii)* health facility service areas can be accurately delineated as a result of the widespread implementation of community-based health insurance.

Population and general spatial data. Demographic data were extracted from the 2002 and 2012 Population and Housing Census published by the National Institute of Statistics of Rwanda (NISR). For the years in between, estimated statistics are extracted from the Integrated Household Living Conditions Surveys (EICV 2 & EICV 3) also reported by NISR. General spatial data (*i.e.* lakes, islands, parks, roads) were obtained from the Centre for GIS and Remote Sensing of the University of Rwanda (CGIS-UR).

Potential associated environmental factors. Several parameters were considered as potential risk factors for STH incidence both at the more general and more detailed spatial scale. The parameters are summarized in Table 10. Three topographical parameters were generated from a high-resolution digital terrain model, produced from 2008 aerial photography (Swedesurvey, 2010). The DEM was used to generate two terrain derivatives, namely slope and terrain shape index (McNab, 1989). Soil parameters were extracted from the soil geo-database of Rwanda, the outcome of a semi-detailed soil survey of 1833 soil profiles spread over the country (Verdoodt & van Ranst, 2006). The available soil properties are pH, and clay and sand percentage.

The annual average of climatic data (temperature, rainfall), measured at 183 weather stations from 1950 to 2010, was interpolated using the thin-plate smoothing spline algorithm as proposed by Hijmans et al. (2005). The boundaries of lakes, wetlands, and main rivers networks, produced by the Rwanda Environment Management Authority (2008) and further adjusted by Nyandwi et al.(2016), were available for this study. For each HFSA, the following proxy factors were generated: (i) area and percentage of wet area (wetland and lakes) per HFSA, (ii) wetland (use) status in terms of the

percentage of area covered by natural vegetation, (Lozano et al.) the total area of wetlands used for intensive (irrigated) agriculture, mostly rice cropping.

Socio-economic conditions per administrative sector were extracted from the 2012 Census (http://statistics.gov.rw/publications). In short, the following socio-economic parameters were considered: *(i)* proportion of rural to urban area, *(ii)* population density, *(Lozano et al.)* percentage of households with (un-) improved water source, *(iv)* percentage of households with access to (un-) improved sanitation, *(vi)* education level (proportion of persons with six years basic education and people with twelve years basic education).

The parameter datasets were obtained from various government organizations: Rwanda Natural Resource Authority, Land, and Mapping department (RNRA/L&M), Ministry of Agriculture and Animal Husbandry (MINAGRI), Rwanda Environmental Management Authority (REMA), and the National Institute of Statistics of Rwanda (NISR).

Table 10. Environmental risk factors used to assess their relationship with STH incidence

Data type	Variable	Source	Description
Topographic	Elevation, Slope, and TSI	RNRA L&M	Using the DEM an elevation map was created and further used to generated Slope and Terrain Shape Index (Hurlimann et al.) using ArcGIS. The outputs have a high spatial resolution with a cell size of 10m.
Climatic	Rain, Temperature Humidity, Evapotranspiration	NMS	Data are measured nationwide at 183 meteorological and agro-meteorological weather stations, data were interpolated and resampled to a cell size of 800m.
Soil characteristics	pH, Clay, and Sand percentage	MINAGRI	Soil geo-database of Rwanda created in 2000 by MINAGRI in collaboration with Gent University/Belgium, Resolution: 500m (1/50 000)
	Wetland area, Wetland proportion, Wetland use	REMA & RNRA	National inventory of wetlands by REMA in 2008 and updated in collaboration with RNRA in 2012; Ortho-photographs resolution: 0.25m
	Rice cropped area	MINAGRI	GIS database of Rural Sector Support Program of MINAGRI, considering new wetland management or rehabilitation for intensive cropping 200-20120.25m
	Demography: Population, Pop. density, Number of Households Residential area Sanitation Water source Education level		

Data analysis
In this study, STH incidence and the relation with correlated factors are assessed at the District and the HFSA level.

Spatial pattern detection and cartographic display of STH incidence rates. The total number of confirmed cases of STH infection per HFSA was aggregated to the District level for 2007 and 2008. The demographic and socioeconomic information were also linked to HFSAs (n=367) and subsequently aggregated to district level (n=30).
For this study, the incidence rate was defined as the number of confirmed STH cases divided by the population (over a time period of 1 year) per HFSA/District, as shown by equation 1.

$$Di = (In / Pt) \times 1\,000\ (0)^{*} \qquad \textbf{\textit{(Eq. 8)}}$$

Where **Di** is the STH incidence rate, **In** is the total number of new cases in 12 months of a year per district or HFSA and **Pt** is the population at risk of that year for a given spatial unit. * Rates are reported per 10.000 persons at the District level, and for 1.000 persons at HFSA level.

Analysis of the spatial distribution of STH incidence was in two stages. In stage *one,* the analysis is done at the aggregate level of the District. District level analysis was done to allow comparison with the outcomes of the nationwide school-based prevalence surveys described in the introduction. Stage *two* analyses STH transmission levels at HFSA level to identify the spatial pattern of STH transmission at a geographically much finer scale (Clements, Deville, et al., 2010).

Spatial autocorrelation analysis and spatial clustering test. Spatial autocorrelation was computed to ascertain if spatial clustering occurs between neighboring HFSAs. For this study, we used Moran's Index test, widely used in the analysis of geographic differences in public health studies (Getis & Ord, 1992). Moran's I was used to identify spatial clusters using Cluster and Outlier Analysis. Negative values indicate negative spatial autocorrelation, positive values indicate the reverse. Values range from −1 (indicating perfect dispersion) to +1 (perfect correlation). A zero value indicates a random spatial pattern (Getis, 2007).

Comparison of the spatial distribution of STHs with incidence and prevalence datasets. Incidence and prevalence are both measures of disease frequency, the 2007/2008 prevalence of each STH type and combined STH infections reported at district level were compared with incidence rates for the same year and district. The prevalence and incidence rates were also mapped at the district level to enable visual comparison.

Assessment of relationships between STHs incidences and potential explanatory variables. Incidence data and potentially associated environmental factors were integrated at HFSA level using ArcGIS 10.2.2. All bio-physical factors maps were standardized to a resolution of 10x10m. Average values for potential explanatory variables for each HFSA and District were extracted using the zonal statistic module of ArcGIS 10.2.2. After that, all data were exported to IBM SPSS Statistics, version 20 for further analysis.

The choice of explanatory variables is an important step when constructing any statistical model (Hessami, Gachon, Ouarda, & St-Hilaire, 2008). Before training the model, the selected variables were screened for collinearity (Ortega Hinojosa et al., 2014). The next step was to split the available HFSA level data (n= 367) into calibration and validation subsets. We randomly selected 60% of the available dataset as training dataset (n= 220) for model fitting. The remaining 40% for model validation (n= 147). The proportion of both subsets followed the principle of using more data for calibration, and a validation subset reaching at least 35% (Lehmann, 1998; Van Oost et al., 2005). Selection of validation and calibration data was done using an online list randomizer tool (https://www.random.org/lists/).

The incidence rates of STHs, as dependent variable y were related to the potentially explanatory variables, as variables X_1 to X_n (soil, topographic, climatic, wetland and socio-economic characteristics). This was done to quantify the strength of the relationship between STHs and the significant covariates (Kabo et al., 2015; Wang et al., 2015). We also report the standardized coefficient to identify which of the independent variables have a greater effect on the dependent variable. This is needed since our variables have different units of measurement (Lemeshow & Moeschberger, 2005).

4.3 Results

Spatial distribution of STH incidence at district and HFSA levels

The total number of confirmed STH infections was 1.204.330 during the period 2007-2008. At the District level, the number of reported cases ranged from 22.000 to 78.000. If we consider individual STH types, very significant differences can be observed (see Figure 17). Rwandans were frequently infected with *Trichuris trichiura* amounting to 57% of all recorded cases, followed by *Ascaris lumbricoides* (38%), and a much lower infection rate of Hookworm (5%). The ratio of the total number of reported cases of individual STH types and combined incidences for the population at risk is visualized in Figure 18a - 18d.

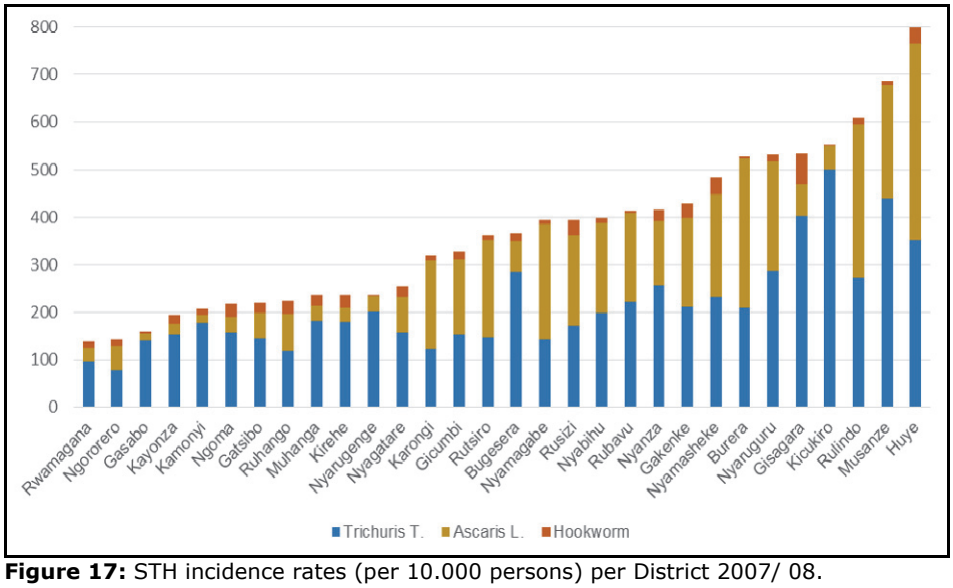

Figure 17: STH incidence rates (per 10.000 persons) per District 2007/ 08.

Incidence rates, here defined as the number of reported cases per 10,000 persons, varies considerably between Districts (see Figure 17). Some Districts (e.g. Rwamagana and Ngororero) have relatively low incidence rates. Others such as, for example, Musanze and Huye overall exhibit high incidence rates.

In Figure 18 incidence rates are classified into low, moderate and high rates using a natural break classification. Combined STH infections show incidence rates ranging between 140 and 800 per 10.000 persons (Figure 18a). *Trichuris trichiura* is a nationwide problem in Rwanda, with incidence rates ranging between 78 and 500 per 10.000 persons (Figure 18b). *Ascaris lumbricoides* incidence follows with rates ranging between 15 and 410 per 10.000 persons (Figure 18c). Hookworm incidence rates are much lower, varying between 3 and 65 per 10.000 persons (Figure 18d). The Northern Province is the most infected region. Kicukiro district (of Kigali city) is one of the Districts with a high incidence rate in general and specifically for *Trichuris trichiura* (500 per 10.000 persons).

Comparing District level prevalence and incidence data

To check if there is a significant relationship between prevalence and incidence data, we compared district level incidence rates of 2007/2008 with district level prevalence data based on the 2007/2008 nationwide school-based survey. Figure 18 illustrates both incidence and prevalence data at District level for individual STH types and for combined infections, as well as scatterplots that illustrate the correlation between the two. We can observe that the correlation between the two data sources varies considerably with a correlation coefficient of 0.63, 0.27, 0.1, and 0.02 for *Ascaris lumbricoides,*

Hookworm, combined STH infection, and *Trichuris trichiura* respectively. The high correlation coefficient for *Ascaris lumbricoides* indicates that routinely recorded cases of *Ascaris lumbricoides* recorded at health facility level can be used to complement prevalence data extrapolated to the District level. To a much lesser degree, this also is the case for Hookworm. For combined STH infections and *Trichuris trichiura*, prevalence and incidence based data are very different.

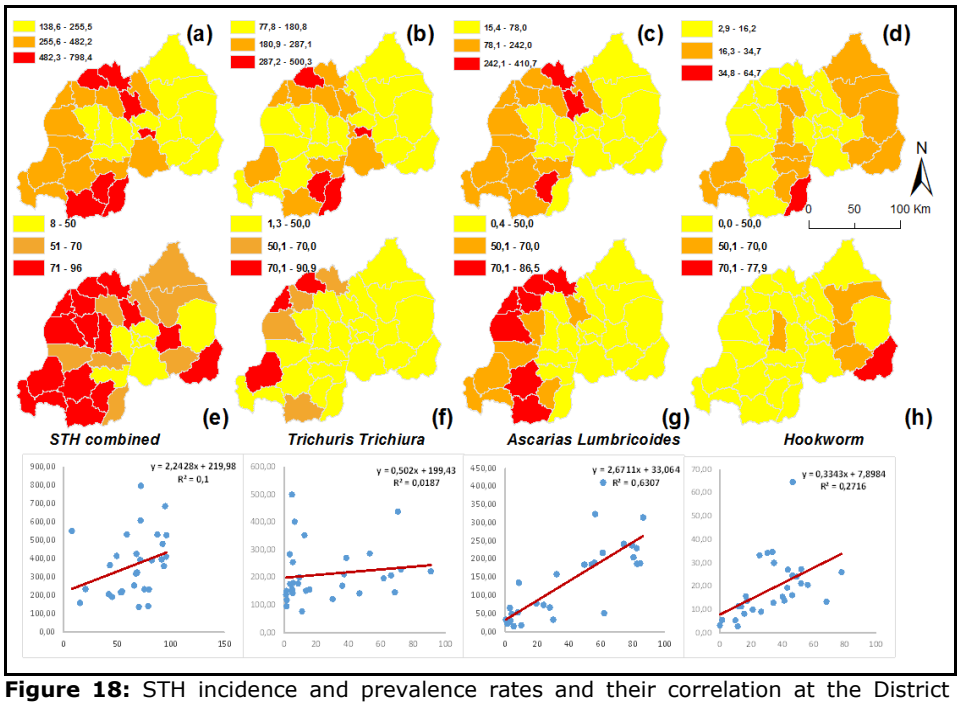

Figure 18: STH incidence and prevalence rates and their correlation at the District level. Starting from the left-upper corner there are maps for combined STHs (**a**), for Trichuris trichiura (**b**); for Ascaris lumbricoides (**c**) and Hookworm (**d**). The prevalence maps and scatter plots are shown for combined STH cases (**e**), individual cases with Trichuris Trichiura (**f**), Ascaris lumbricoides (**g**), and Hookworm (**h**).

Spatial patterns and clusters of STH transmission at the more disaggregated HFSA level.

In the following, we further analyze the spatial pattern and clustering of STH transmission at the detailed spatial scale of the HFSA using routinely recorded data of confirmed cases, as illustrated in Figure 19.

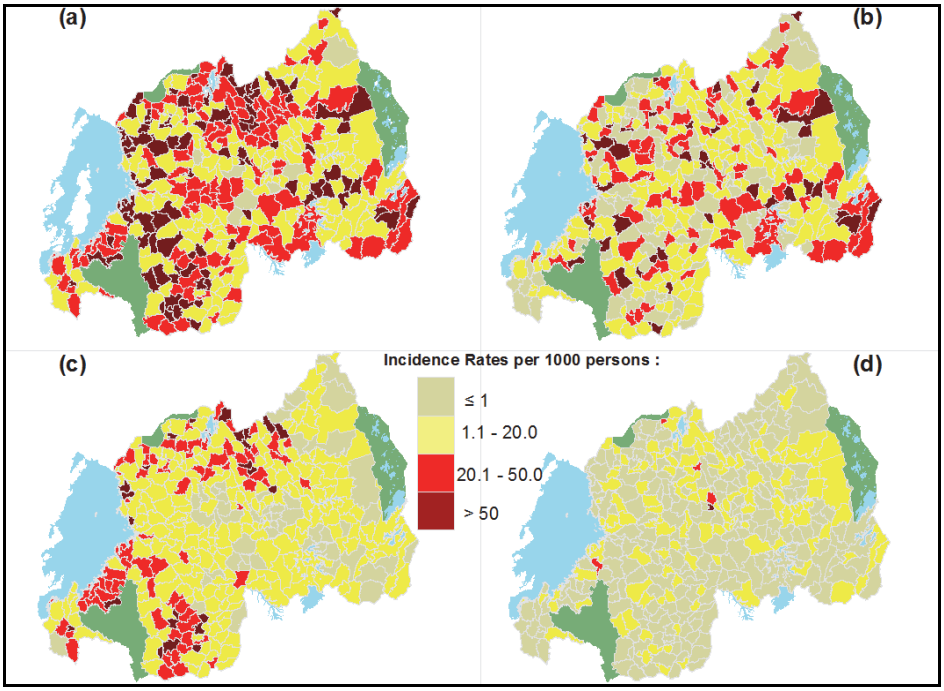

Figure 19: Incidence rates at HFSA level. The rates are presented for combined STH infection (**a**), Trichuris trichiura (**b**), Ascaris lumbricoides (**c**) and Hookworm (**d**).

Figure 19a shows that combined STH infection is widespread in Rwanda with medium to high-level incidence rates in many HFSAs throughout the country. The spatial pattern of *Trichuris trichiura* is overall similar as visualized in Figure 19.b. *Ascaris lumbricoide*, on the other hand, exhibits a much more focalized spatial distribution with areas of high transmission in the northern, southern and south-western areas of the country (see Figure 19.c). As illustrated in Figure 19.d, only a few HFSAs have high transmission rates of Hookworm (namely Kabuga HFSA in Kamonyi district, Gatare HFSA in Nyamasheke district, Rwahi HFSA in Rulindo district, Rutenderi HFSA in Gakenke district, and Mucaca HFSA in Burera district). Apart from these areas, Hookworm does not have high incidence rates in Rwanda.

Spatial auto-correlation and cluster detection. The spatial auto-correlation of incidence rates per STH type at the detailed HFSA level was further confirmed by a positive and statistically significant spatial autocorrelation as illustrated by the result of computed Moran's Index in Table 11.

Table 11. Moran's Index calculated for each STH and combined incidence rates at HFSA level.

Year	Moran's I	P-value	z-score
Combined STH	0.113	0.000006	4.51
Trichuris trichiura	0.052	0.037	2.08
Ascaris lumbricoides	0.375	0.000000	14.433
Hookworm	0.0056	0.744	0.326

Moran's I statistic indicates that there is a less than 1% likelihood that the clustered pattern could be the result of random chance for *Ascaris lumbricoides* and for combined STH infections. Spatial clustering of *Trichuris trichiura* has a less than 5% likelihood to be the result of random chance. The spatial pattern of Hookworm is random.

The Anselin Local Moran's I cluster analysis results at HFSA level has three forms of spatial association (see Figure 20): not significant association, the positive spatial association observed (in the area marked with High-High and Low-Low clusters), and negative spatial association characterized by dissimilar values (High-Low and Low-High outlier).

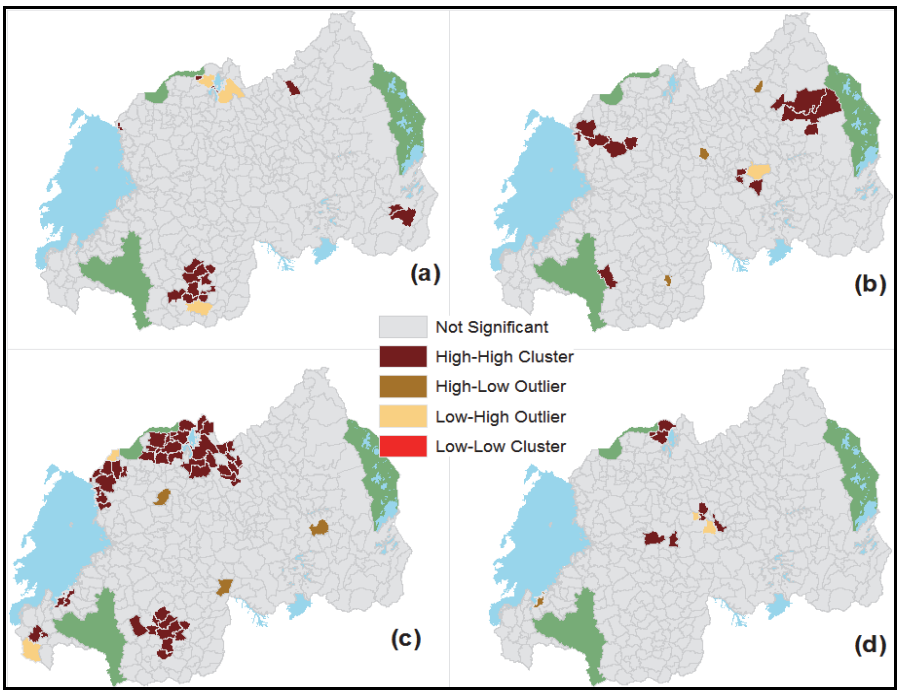

Figure 20: Spatial distribution of Anselin Local Moran's I values at the HFSA level. The index for (**a**) Combined STH cases; (**b**) *Trichuris trichiura*; (**c**) *Ascaris lumbricoides* and (**d**) Hookworm.

STH incidence at District and HFSA level and their association with environmental factors.

The best model was obtained using a stepwise linear regression with a random distribution of residuals producing the highest R-squared and lowest standard error. The environmental factors associated with STH incidences vary per STH type and with the spatial scale of analysis. Table 4 summarizes statistical outputs of correlated ecological factors (related to wetland environment: low elevation, less sandy soil, wetland covered area and their uses), climatic conditions (rainfall), and socio-demographic characteristics (rural population, the number of households, poor sanitation condition).

Table 12. Standardized Beta Coefficients of associated factors at District and HFSA spatial scale

Factors	*Combined*		*Trichuris. Tr.*		*Ascaris. L.*		Hookworm	
	District	*HFSA*	*District*	*HFSA*	*District*	*HFSA*	*District*	*HFSA*
pH	-	-	0.43**	-	-	-	-	-
Sand prop (%)	-0.38**	-0.4*	-	-	-	-0.4**	-	-
Elevation (m)	-0.53**	-	-	-	-	-	-	-0.29**
Wetland area (ha)	-	-	-	-	-	-	-	-
Wetland proportion (%)	0.39**	0.22*	-	0.5**	-	-	-	-
Wetland cropped -Rice (ha)	-	-	-	0.26*	-	-	-	-
Rainfall	0.43**	0.29*	-	-	-	0.3**	-	-
Number of HH	-	-	-	-	0.37*	-	-	-
Unimproved sanitation	-	-	0.37*	-	-	-	-	-
Rural	-	0.13*	-	-	-	0.12*	-	-
Urban	-	-	-	-0.27*	-	-	-	-

Coefficient presented are significant with: $p < 0.05$ (*) and $p < 0.01$ (**)

The exploratory regression models indicate that environmental factors explain between 13 - 30% of total variability in *Trichuris trichiura*, *Ascaris lumbricoide*s and combined STH cases at different spatial scale. The combined STH incidences model reached 82% at the District level. Hookworm incidence was poorly explained by available risk factors, reaching only 8% at HFSA level (see Table 13).

Table 13. Summary of linear regression models for incidence rates for STH

Spatial level	Factor	Significant Variables	R^2	Std. Error of the Estimate
Ascaris lumbricoides				
	Climatic	Rain		
	Physical	Sand percentage		
	Socio-economic	Rural prop		
District	Demographic	Number of HH	0.13	2952.698
Hookworm				
HFSA	Physical	Elev.	0.08	52.635
District				
Trichuris trichiura				
		Wetland cultivated area		
		Wetland proportion		
	Demographic	Urban proportion		
	Physical	pH		
	Demographic	Unimproved sanitation		
Combined STH				
	Climatic	Rain		
	Physical	Sand percentage		
	Ecological	Wetland prop		
	Demographic	Rural prop		
	Physical	Elev., Sand percentage		
	Climatic	Rain		
	Ecological	Wetland area		

4.4 Discussion

Findings illustrate that the spatial distribution of combined STH incidence is quite different compared to individual STH types. This clearly indicates that it is not very meaningful to report a combined STH pattern. The readily available systematically recorded routine health data at detailed spatial scale makes the spatial assessment of individual STHs simple and cheaper in comparison to conducting prevalence studies. By using data of individual STH types, we can separately assess the relationship between potentially associated risk factors and STH incidence more directly. Although STHs are considered to be widespread in Rwanda, the distinctive local characteristics

and the association with wetlands and wetland use is significant. The clearly different spatial patterns identified and the differences in the underlying risk factors can support STH specific intervention campaigns instead of the current mass drug administration approach.

Incidence rates versus prevalence data at the District level. We observed a clear relationship between the incidence and prevalence data at the District level. There is a significant positive correlation between incidence and prevalence data for *Ascaris lumbricoides* and hookworm. While for *Trichuris trichiura* and the STH combined only a weak positive significant correlation was observed. Prevalence data are extrapolated to the District level on the basis of results obtained at 4 to 6 surveyed schools per District. The Rwandan cross-sectional prevalence survey follows the World Health Organization guidelines (Montresor et al., 1998), but has some limitations. One limitation is that the sampling strategy is not spatially stratified, and includes only a limited number of schools to represent a large geographical area with internally varying risk of infection (Clements, Deville, et al., 2010). Schur et al. (2013), make use of advanced geostatistical technologies to extrapolate a limited number of field observations to a larger geographical area. In addition, the primary school based prevalence surveys, focus on children as an only exposed group. In reality, the increase of wetland reclamation for rice cultivation during last decade (Malesu et al., 2010) has most probably increased exposure of adults involved in rice cropping activities.

Different spatial patterns for individual STH types. Routinely collected confirmed case data captured at health facility level can be used to adequately identify and visualize the spatial pattern of STH incidence at a fine geographic resolution, as illustrated in Figure 17. STH transmission is not uniformly distributed over the country. The presence of spatial clusters, detected by Moran's I and LISA (Table 11 and Figure 19), further confirms this. Previous studies have also reported on the disparate spatial distribution of different STH types (B. Abera, Alem, & Melaku, 2012). This study has highlighted that *Trichuris trichiura* is geographically widespread in Rwanda, with 57% of the population infected. More importantly, the spatial pattern of the three STH types is very different. This is in line with Clements, Deville, et al. (2010), who also identified different spatial patterns for the same STH types in Uganda, Tanzania, Kenya, and Burundi. There are relatively few cases of Hookworm compared to other intestinal helminths in Rwanda. Since the transmission to humans usually occurs through bare feet on contaminated soil, the limited number of reported cases may be linked to improved socio-economic conditions in Rwanda during the last decade (MINECOFIN, 2015) leading to increased wearing of shoes.

STH type incidence is associated with different environmental variables. The regression models demonstrate different fits at District and HFSA level, and with different environmental conditions. A typical phenomenon with this type of analysis is the observation that model fit increases with aggregation level (Kok and Veldkamp, 2011). Although this is the case for combined STHs, it is not that clear if we consider the individual STH types. The multiple regressions have generated models with four independent contributions from soil sand content, rainfall, area proportion of wetland, and demography. It appears that, in general, the STHs are more likely to infect people in densely populated rural areas with less sandy soil, higher rainfall, and in wetland-rich areas. The differences between associated factors and STH incidence appear to be related to different helminth living conditions and different transmission mechanisms, which vary geographically (Larsen & Roepstorff, 1999). *Ascaris lumbricoides* incidence is mainly associated with densely populated rural areas with less sandy soils and high rainfall. *Trichuris trichiura* incidence, on the other hand, is clearly associated with wetlands outside urban areas, while Hookworm is weakly associated with lower altitudes.

Many of the associated biophysical factors are directly or indirectly related to wetlands environments. In Rwanda, wetlands are seen as the only alternative to increasing food production and ensuring food security for the population. During last decade rice has become a major food crop in Rwanda, the total area of rice production had increased 4-fold from 4.000 to 16.000 hectares between 2000 and 2010 (Malesu et al., 2010).

The relationship between high population density and intestinal helminth incidence comes from residential settings allowing favourable conditions for transmission (Macchioni et al., 2015). In Rwanda, with persistent and growing urban and sub-urban agriculture (Thaxton, 2009), this remains a problem.

Limitations of the study. The number of confirmed cases recorded at health facilities may be only a fraction of all infected persons for a number of reasons. Firstly, self-treatment is common in Rwanda, as it is in most of developing countries (Phillips, 1990). Most families practice regular de-worming using either traditional medicinal treatment or anti-intestinal medication obtained from a pharmacy. Most probably, reported cases will concern patients with severe infections and with clear clinical symptoms. Secondly, routine health data used for this study are collected from public and faith-based health facilities only. Since the proportion of people seeking private health care in Rwanda is limited this is considered as a minor shortcoming only (National Institute of Statistics of Rwanda, 2012). Thirdly, the proportion of reported cases may also be influenced by variations in

health seeking behaviour. This may explain the high number of confirmed STH cases in urbanized areas (i.e. Kicukiro district of Kigali city and Huye district) inhabited by relatively well-educated population groups.
On the positive side, the Rwandan Health Management Information System systematically and accurately records laboratory confirmed STH cases at a fine geographic resolution (USAID/Rwanda, 2006). Because of this, we feel that recorded STH cases adequately describe the spatial and temporal patterns of transmission for individual STH types.

4.5 Conclusion

Routinely collected confirmed STH cases are suitable to generate spatially explicit overviews of the distribution and intensity of STH transmission for small geographic units which we here termed HFSA. The distribution of individual or combined STH incidence in Rwanda varies per helminth type and with environmental conditions. It is also clear that the STH occurrence is geographically clustered in areas much smaller than the District level which has been used to date. *Trichuris trichiura* is highly endemic with a broad spatial distribution throughout the country. Hookworm has the lowest incidence rates with an almost random spatial distribution. A significant geographical association is observed between STH incidence and specific biophysical and socio-demographic factors. The most strongly associated biophysical factors are directly or indirectly related to wetland environments. The number of STH cases is also associated with increasing rainfall, population density, the proportion of the rural population, the area covered by wetlands and rice cropping, less sandy soils and at lowest elevation levels.
Since the routine health data used in this study are readily available, the approach described in this study can be used to complement studies based upon prevalence survey data. Finally, the presented methodology can contribute towards more effective, location specific, STH control programs design and evaluation.

Chapter 5: Spatio-temporal dynamics of schistosomiasis in Rwanda between 2001 and 2012: impact of the national NTD control programme[4]

[4] This chapter is based on: Nyandwi, E., Veldkamp, T., Osei, F. and Amer, S. Spatio-temporal dynamics of schistosomiasis in Rwanda between 2001 and 2012: impact of the national NTD control programme, Geospatial Health 2017, volume 12 (1), 48-58.

Abstract

Schistosomiasis is recognized as a major public health problem in Rwanda. We aimed to identify the spatiotemporal dynamics of its distribution at a fine-scale spatial resolution and to explore the impact of control programme interventions. Incidence data of Schistosoma mansoni infection at 367 health facilities were obtained for the period 2001-2012. Disease cluster analyses were conducted using spatial scan statistics (SaTScan) and geographic information systems (GIS). The impact of control interventions was assessed for three distinct sub-periods. Findings demonstrated persisting, emerging and re-emerging clusters of schistosomiasis infection across space and time. The control programme initially caused an abrupt increase in incidence rates during its implementation phase. However, this was followed by declining and disappearing clusters when the programme was fully in place. The findings presented should contribute to a better understanding of the dynamics of schistosomiasis distribution to be used when implementing future control activities, including prevention and elimination efforts.

Keywords: Spatiotemporal, schistosomiasis, Schistosoma mansoni, cluster, health facility service area, Rwanda.

5.1 Introduction

Reliable updates of the number of people currently infected by schistosomiasis worldwide are difficult to come by as it depends on the level of sensitivity of the diagnostic techniques used. Although the figure of 250 million infected victims given by Hotez (2014) is probably an understatement, it is close to figures previously given by the World Health Organization (WHO) and most authors. Mortality is even more difficult to pin down but the figure provided by the Global Burden of Disease (GBD) study in 2010 (Lozano *et al*. 2012) shows that the disease is primarily chronic. Left untreated, schistosomiasis causes damage to the bladder, kidneys, liver, spleen and also other organs depending on which species of the parasite is involved (van der Werf, 2003).

Schistosomiasis is an important neglected tropical disease (NTD) recognized as a major public health problem (TRAC+, 2008). In Rwanda, intestinal schistosomiasis caused by *Schistosoma mansoni* is the only form of the disease, so the morbidity seen does not include the urogenital system. Until 2007, the spatial prevalence pattern within Rwanda remained unidentified. In that year, the Rwandan Ministry of Health (MoH) established an NTD control program which started with a nation-wide mapping of NTD prevalence based on a cross-sectional survey of randomly sampled primary schools in each of the country's 30 districts (TRAC+, 2008). School survey outcomes were extrapolated to district-level prevalence maps, which were subsequently used to guide a control initiative (Figure 21). The initiative started with a three-year project (2008-2010) with several intervention measures targeting the most endemic districts. The nationwide NTD interventions included mass drugs administration (MDA), training of health workers on diagnosis, treatment and transmission control. Health promotion materials were also distributed.

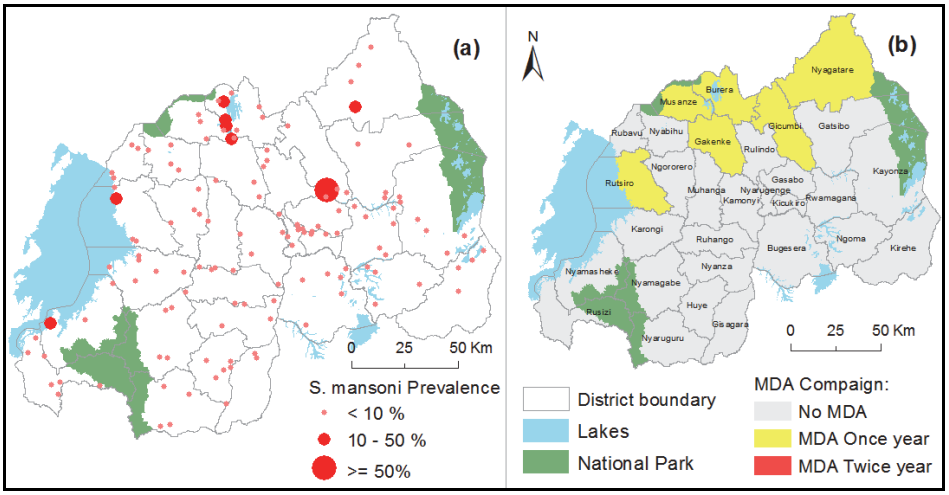

Figure 21: Distribution of Schistosomiasis in Rwanda with an indication of areas prioritized for control. **(a).** Prevalence map from 2007/08 school-based prevalence survey. **(b)**. Areas prioritized for MDA.

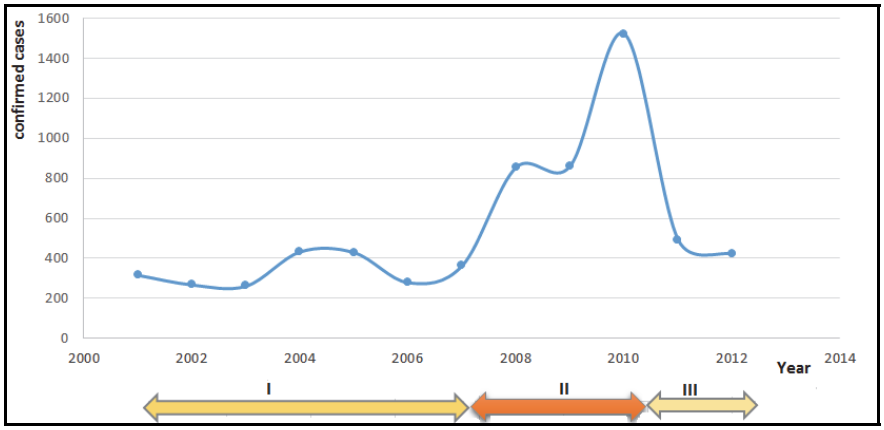

Figure 22: Reported schistosomiasis cases between 2001- 2012. The three identified time periods are I (before the NTD control programme); II (during the three years of NTD control initiation); and III (the period when the NTD program had become fully functional).

Figure 21 shows the schistosomiasis prevalence in the 2007/08 period, including the implementation of MDA using praziquantel, while Figure 22 illustrates the achievement of the control initiative. We considered three subsequent periods of the control programme, the first of which shows the period before its establishment (up to in 2007), the second, the implementation period that was surprisingly associated with an abrupt increase (2008 – 2010). However, this increase was most probably attributed to improved diagnosis that was part of the programme implementation. The third period from 2010 onwards depicts the programme when fully operational, which was characterized by a steep decline in the number of

incidences, apparently stabilizing to a level comparable to the pre-2007 situation. Thus, despite the control efforts, the number of confirmed cases of schistosomiasis remains high at the national level. Recent studies by Ruberanziza et al. (2010, 2015) also identified changes of focal areas with relatively high prevalence. Some of the recently identified areas with high prevalence are located in districts initially classified as being of low endemicity.

Since schistosomiasis can have a highly localized distribution (Gray et al., 2011), a spatiotemporal assessment based on routinely collected incidence data at primary health facilities, can be done using local clustering methods to identify areas with high disease occurrence (Anselin, 1995; Song & Kulldorff, 2003; Bernasco & Elffers, 2010; Quick & Law, 2013). Local clustering methods can be used to identify if cases are geographically concentrated (spatial clusters), tend to be placed closer in time (temporal clusters), or are close both in space and time - spatiotemporal clusters (Tango, 2010). Few studies, however, have aimed at the detection of spatiotemporal schistosomiasis at a detailed spatial scale. A notable exception is Gao et al. (2014) who analyzed spatiotemporal clustering of schistosomiasis japonica at the village level in Anhui province in China.

The objectives of this study were to explore the spatiotemporal distribution dynamics of schistosomiasis in Rwanda between 2001 and 2012 and to investigate the linkage of this distribution with the NTD control programme interventions. We aimed to analyze these outcomes to assess the impact and effectiveness of the current elimination strategy in Rwanda.

5.2 Materials and Methods

The systematic spatiotemporal variation of *Schistosomiasis mansoni* incidence rates at HFSA was detected with Kulldorff's Spatial Scan Statistics (SaTScan v9.4.2) and results were visualized with GIS software ArcGIS, v. 10.4 ESRI, Redlands, CA, USA).

Study area

Data on schistosomiasis infection and population

Patients suspected to have schistosomiasis are asked and instructed how to collect and submit stool specimen, which was subsequently subjected to testing for eggs of soil-transmitted helminths (STH) and *Schistosoma mansoni* using the Kato-Katz technique (Katz et al., 1972). Trained medical laboratory technicians conducted the sample analysis. Two slides per stool specimen were prepared and read separately. The result recorded per patient was the mean of the reading of the two slides (egg count or a number of

eggs per gram of stool). The number of confirmed schistosomiasis cases per primary health facility were recorded and first reported in an Excel table and later loaded into the Rwandan Health Management Information System (R-HMIS) every month by each primary health facility using a web-based software platform called DHIS 2 (USAID/Rwanda, 2006). The recorded cases, without personal information, were provided by the Malaria and Other Parasitic Diseases Division of the Rwanda Biomedical Centre (RBC/M&OPD). The dataset consisted of laboratory confirmed cases from all public and faith-based primary health facilities for the period January 2001 to December 2012.

The disease data were felt to be representative for several reasons, such as (i) Health service accessibility - primary health facilities are within walking distance for most people (<5 km distance); (ii) Affordability - health care for most Rwandans is facilitated by the Community-Based Health Insurance - patients pay only 10 % of the total cost of service and medication (Lozano et al., 2012), Encouragement - community health workers actively stimulate patients to visit the primary health facility in case of suspected health problems and (iv) Medication - there is no traditional medicine used for schistosomiasis in Rwanda, patients with symptoms are treated at the primary health facility.

Demographic data were extracted from the 2002 and 2012 Population and Housing Census published by the National Institute of Statistics of Rwanda (NISR). For the years between, population data were estimated using the average national population growth rate of 2,6 % (NISR, 2014). The resulting population data from 2001 to 2012 for each HFSA was prepared as input for the space-time analysis in SaTScan. In addition, the coordinates file was prepared, as X, Y of centroids for each HFSA generated in ArcGIS 10.4.

Spatial and space-time clustering analysis

The annual number of confirmed schistosomiasis cases per HFSA were stored in Excel tables. Kulldorff's spatial scan statistic (SaTScan), a wide application in epidemiology (Hanson & Wieczorek, 2002; Read et al., 2011; Yao et al., 2011), was used to identify HFSAs with significantly high incidence rates. Given that we had data on confirmed cases for the period 2001-2012 a retrospective analysis (spatial and space-time) was applied. The analysis was performed using SaTScan 9.4.2 (Kulldorff, 2015).

The advantage of the SaTScan over other local cluster analysis methods is that it calculates a likelihood test statistic, avoids multiple testing and has a variable scan window which prevents pre-selection bias (Mather et al., 2006). SaTScan determines statistical significance without specifying the number of areas or location of the clusters before calculating the level of significance; all

significant clusters must reject the null hypothesis based on their strength (Almeida, et al., 2011; Kulldorff, 2015). Because of this, any rearrangement of cases outside of the scan window will not change the cluster significance level (Song & Kulldorff, 2003). The level of significance is determined through Monte Carlo hypothesis testing, an estimate of the rank of the likelihood of the real data based on randomized versions of the same dataset (Kulldorff, 2015). The likelihood ratio is calculated as:

$$LR(Z) = \frac{L(Z)}{L_0} = \left(\frac{c(Z)}{n(Z)}\right)^{c(Z)} \left(\frac{C - c(Z)}{C - n(Z)}\right)^{C - c(Z)}$$

(Eq 9)

where *LR(Z)* represents the likelihood ratio for scan window *Z* which is calculated as the likelihood of the alternative hypothesis of spatial clustering (*L(Z)*) divided by the likelihood of the null hypothesis of complete spatial randomness (L_0). *c(Z)* is the observed number of cases in scan window *Z*, *n(Z)* the expected number of cases inside *Z*, and *C* the total number of cases. After the likelihood ratio (LLR), a *p*-value is estimated by comparing the rank of the maximum LLR from the real data set with the highest LLR from the random datasets using Monte Carlo simulation. We set the number of replications to 999 times. Clusters with *p*-values <0.05 indicate raised risk. Relative risk (RR) was calculated for each statistically significant cluster (Kulldorff, et al., 1998) by comparing the risk within the cluster in a particular time and area with the risk outside the cluster.

Spatial clustering analysis

Spatial clusters were detected using SaTScan's circular window that scans the entire study area. The radius of the circle can be varied continuously from zero to a specified maximum spatial window size, corresponding to the maximum percentage of the total population at risk. Since we expected to identify highly localized clusters, we followed the recommendation of Gao et al. (2014) and opted for the maximum window size to contain 10% of the population at risk. For each window, the observed cases inside and outside the circle were compared to the number of expected cases, as calculated using the Poisson distribution. The window with the maximum likelihood ratio was identified as the most likely (primary) cluster (Kulldorff, 1997). Subsequent likelihood ratios with significance levels of < 0.05 were considered secondary clusters.

Space-time clustering analysis

For detection of space-time clusters, the cylindrical approach was used. Here the window was used as described above, while the temporal aspect was expressed as the height of a cylinder with that window as a base. The window started with the minimum radius and height at one location and was

then moved throughout the study area with a continuously varying radius and height until reaching the upper limit of the radius. The maximum spatial window was set at 10% of the population at risk, while the height was set at 50 % of the total length of the study period.

Cluster detection within sub-periods of study period

Figure 22, identifies three sub-periods that coincide with the time before, during and after the implementation of the NTD control programme. Therefore, the methodology described above was first applied to the entire study period and then repeated for each of the three identified sub-periods. The underlying motivation for this was to facilitate the analysis of the impact of the intervention.

5.3 Results

Investigation of spatial schistosomiasis clusters

Transmission of schistosomiasis was found to be concentrated in 32 spatial clusters as shown in Table 14. The most likely cluster found was Ntaruka HFSA situated in Burera District in the Northern Province. The secondary most likely cluster was seen in Jarama HFSA located in Ngoma District in the Eastern Province. The largest cluster consisted of 18 HFSAs located in Rusizi and Nyamasheke Districts in the Western Province.

Table 14. Spatial clusters of schistosomiasis 2001- 2012

Cluster	HFSAs involved	Name(s) of HFSAs with clusters	LLR[a]	RR[b]	Obs[c]	Exp[d]
1	1	Ntaruka/Kinoni	2828	197	668	4
2	1	Jarama	807	26	356	14
3	2	Cyanika, Kinyababa	597	10	438	47
4	18	Bushenge, Bugarama, Gihundwe, Gisakura, Kibogora, Mashesha, Mibilizi, Nyamasheke, Mushaka, Muyange, Mwezi, Nkombo, Nkungu, Nyabitimbo, Mukoma, Rangiro, Ruharambuga, Rusizi	576	4	1156	374
5	1	Bugarama	517	9	403	47
6	5	Cyabayaga, Karama, Mimuli, Rukomo, Rurenge.	443	7	444	73
7	1	Nkombo	432	20	212	11
8	4	Cyondo, Mimuli, Muhambo, Rukomo.	354	7	327	48
9	1	Kibogora	254	11	176	17
10	1	Rwesero	200	27	86	3
11	1	Rusoro	188	10	132	13
12	4	Cyabingo, Murandi, Nyundo, Rwaza.	181	5	236	51
13	1	Cyimbili	148	10	109	12
14	1	Gikonko	130	10	96	10
15	1	Mwendo	121	14	72	5
16	1	Bisate	85	6	87	14
17	1	Rugarama	80	4	117	28
18	2	Murara, Byahi/Gacuba.	79	3	168	53
19	1	Nzige	74	6	84	15
20	2	Ntaruka/Nasho, Gituku	64	7	62	9
21	3	Gasiza, Kinoni-Gahunga, Mucaca.	61	3	143	48
22	1	Humure	54	7	53	8
23	4	Gisakura, Mukoma, Muyange, Nyamasheke.	49	3	141	54
24	2	Nyagahita-Nyagatare, Rurenge.	45	4	83	24
25	1	Busanza	37	3	101	38
26		Mashesha	31	4	47	11
27	1	Rusizi	19	3	61	25
28	4	Mukono, Bushara, Bwisige, Cyondo	18	2	77	35
29	1	Nyamyumba	15	4	25	6
30	1	Nyacyonga	13	3	36	14
31	1	Gatare-Mugonero	12	4	21	6
32	1	Kigufi	9	2	41	19

[a] Log Likelihood Ratio; [b]Relative risk; [c] Observed number of cases in the cluster; [d] Expected number of cases in the cluster; All clusters had statistical significance levels (p- value) < 0.01.

Investigation of schistosomiasis clusters

As shown in Figure 23 and Table 15, sixteen spatiotemporal clusters were identified between 2001 and 2012. The most likely cluster had a high LLR of 2798,3, and the least likely had an LLR of 28,4 (p<0,001). This

demonstrated that a statistically significant clustering pattern of schistosomiasis incidence existed both in time and space.

The most likely cluster was Ntaruka HFSA showing high rates for the six-year period of 2003-2008. A similar persistence in time was observed in cluster 4 (2007-2012) and cluster 10 (2005-2010). Many spatiotemporal clusters existed only for one year (in 2008, 2009 and 2010).

If we consider geographical distribution, all 16 spatiotemporal clusters were situated near lakes, main rivers and floodplain wetlands, which all are known for schistosomiasis transmission due to the presence of habitats of the intermediate snail host.

Table 15. Spatio-temporal schistosomiasis clusters in order of LLR

Cluster	HFSAs involved	Timeframe	Name(s) of HFSAs with clusters	LLR[a]	RR[b]	Obs[c]	Exp[d]
1	1	2003 - 2008	Ntaruka	2798	348	582	2
2	1	2010	Jarama	1631	283	353	1
3	9	2008 - 2010	Cyabayaga, Cyondo, Karama, Muhambo, Mimuli, Nyarurema, Rukomo, Rurenge, Tabagwe.	763	14	451	34
4	18	2007 - 2012	Bugarama, Bushenge, Gihundwe, Gisakura, Kibogora, Mashesha, Mukoma, Mibilizi, Mushaka, Muyange, Mwezi, Nkombo, Nkungu, Nyabitimbo, Nyamasheke, Rangiro, Ruharambuga, Rusizi.	639	5	869	202
5	1	2010	Rusoro	428	102	118	1
6	18	2008 - 2010	Busanza, Gikomero, Gishali, Giti, Humure, Kayanga, Kinyinya, Masaka, Muhura, Murambi-Musha-Gahengeri, Muyumbu, Nyagasambu, Nzige, Rubungo, Ruhunda, Rwesero, Rwimitereri, Tanda,	2798	348	582	2
7	1	2010 - 2012	Mwendo	1631	283	353	1
8	4	2010	Byahi-Gacuba II, Gisenyi, Kigufi, Murara.	763	14	451	34
9	1	2009 - 2011	Gikonko	639	5	869	202
10	1	2005 - 2010	Cyimbili	428	102	118	1
11	1	2009	Nyacyonga	336	5	438	95
12	1	2010	Nyamyumba	205	48	71	1
13	4	2009	Gisovu, Mugonero, Mukungu-Bisesero, Nyarwungo.	198	15	111	7
14	3	2008	Jali, Kabuga-Kamonyi, Rutonde.	159	25	70	3
15	1	2008	Nyamure	156	15	88	6
16	2	2008	Mbazi and Sovu	87	30	36	1

[a]Log Likelihood Ratio; [b]Relative risk; [c]Observed number of cases in the cluster; [d]Expected number of cases in the cluster; All clusters had statistical significance levels (p- value) < 0.01.

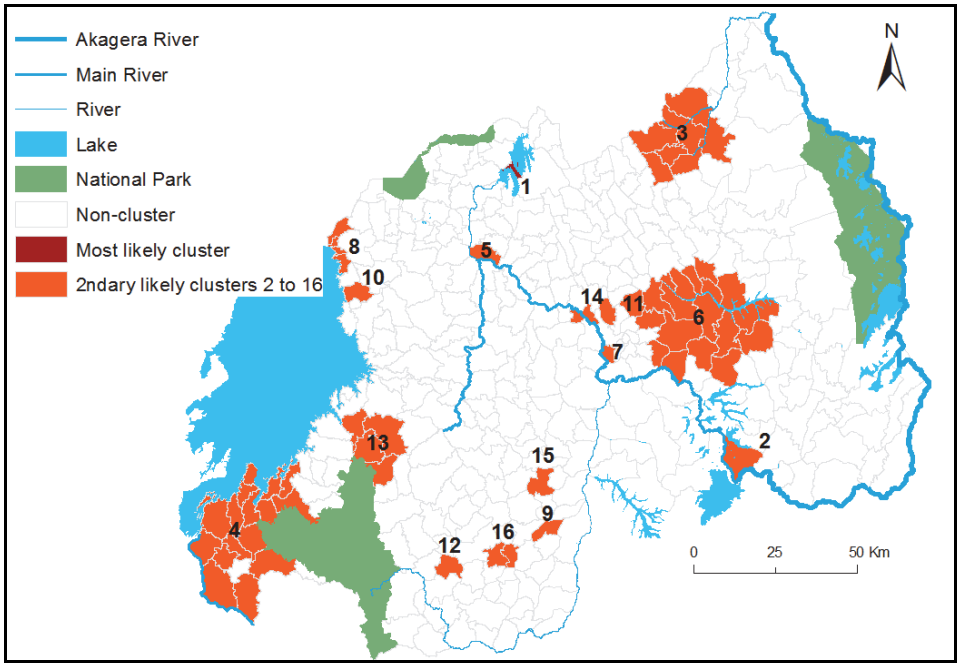

Figure 23: Spatiotemporal schistosomiasis clusters in the period 2001-2012

Relation to the NTD control programme

Clustering in each sub-period

The results of the spatial scan statistic for each of the three sub-periods is visualized in Figure 24. There were 19, 29 and 11 clusters observed in Rwanda for sub-periods I, II, and III, respectively. While the LLR of the primary cluster remained very higher compared to the LLR value of the secondary likely cluster 2, the geographical location and size of the identified most likely clusters differed for each of the three sub-periods.

During sub-period I (Figure 24a), the most significant schistosomiasis cluster was in the North, corresponding to the Ntaruka HFSA. During sub-period II the primary cluster was in the Southeast, i.e. the Jarame HFSA located in Ngoma district, Eastern Province (Figure 24b). In sub-period III, the most likely cluster, covered very a large area, in the South-West (see Figure 24c). In sub-period III all clusters were located in the periphery of the country, with the central parts no longer exhibiting high rates of infection. Cluster number 4 (Kigali City) and 6 (Gisagara District) were, however, exceptions. This meant that hotspots had shifted to peripheral areas, with the most likely cluster in the extreme Southwest and numerous new hotspots in the eastern periphery.

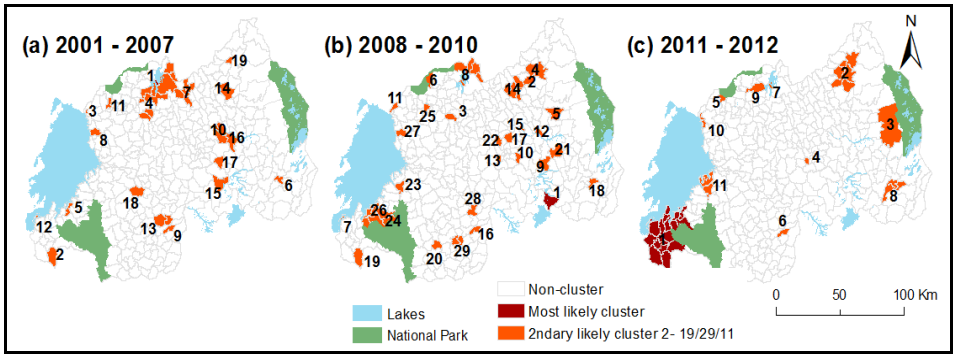

Figure 24: Spatio-temporal schistosomiasis clusters for the three sub-periods. (a) Sub-period I (before NTD control program creation); **(b)**, sub-period II (during the first NTD control intervention); and (**c**) sub-period III (with the NTD control programme in place.

Spatial and temporal cluster dynamics

We distinguish three categories of clusters. The first category consisted of persistent clusters, i.e. areas characterized by high rates of schistosomiasis in each of the three sub-periods. There are four such persistent areas (Nkombo Island, Bugarama, Ntaruka-Kinoni, and Gikonko) that were seen in the whole study period, and appeared in each sub-period. The second category were emergent clusters, characterized by the presence in areas which originally are not highly prevalent but became so after 2007. The third category consisted of disappearing clusters, areas with high rates of schistosomiasis during sub-period I but not seen during the two later in sub-periods. In these clusters, the rate declined significantly over time.

As can be seen in Figure 21a&b, the prioritized areas target zones with prevalence greater than 10%. The mass MDA was planned at the district level using prevalence classes as follow: no MDA for district with <10% prevalence; once a year for school children and women of a child-bearing age where prevalence was < 30% and twice a year all where the prevalence was greater >30%.

5.4 Discussion

The focal spatial distribution of *schistosomiasis mansoni* in Rwanda was confirmed by the spatial scan statistics space-time analysis used in this study methods based upon. As illustrated in Figure 22, the annual number of confirmed schistosomiasis cases in the period 2001-2012 exhibited considerable variation over time. During the first seven years, the period before the NTD control programme, we observe slightly varying medium-level numbers of infections that increased sharply during the first three years of control implementation. At first glance, this may come as a surprise, but the explanation is surely a reflection of improved diagnostic coverage. From 2010

onwards, we see a steep decrease in annual incidence, stabilizing at the pre-intervention period, which shows the impact of the NTD intervention programme (as the diagnostic coverage remains the same). Taking the annual population growth of 2.6% into account, we can conclude that incidence rates have been dynamic during the study period. Some HFSA were persistently classified as schistosomiasis cluster-related with different levels of significance. Over time, new focal areas emerged, while others disappeared. Overall, 67 HFSA units out of a total of 367 were identified as areas with increased transmission of the disease. Persistent and emergent clusters amounted to 53 HFSA units, while 14 such units became non-endemic towards the end of the study period.

The observed spatiotemporal changes in high-rate schistosomiasis clusters were associated both with changes in environmental risk factors and impacts of the various interventions measures initiated by the NTD programme. In the following, we discuss in more detail the association between these clusters and the interventions. Furthermore, the observed spatiotemporal dynamics may be explained by analogy from the literature. We structure this discussion by presenting possible relations for persistent clusters, emergent clusters, and disappearing clusters.

Starting at the first three years of the NTD control program, resources for schistosomiasis control (awareness raising, diagnostic equipment, medication, training of health personnel) increased tremendously (Ruxin & Negin, 2012). In three years, nearly 100% of the population at high risk was covered by at least two MDAs within the priority districts (Sabin Vaccine Institute, 2011). The treatments were accompanied by the distribution of informational and educational materials. At the end of the project, 18 825 health staff were trained to identify symptoms and treatment of intestinal schistosomiasis (Ruxin & Negin, 2012). Increased knowledge of, and means for diagnosis and treatment, probably explains the abrupt increase of confirmed cases during the sub-period II of the program implementation and the general strong decrease of confirmed cases since 2011.

Our spatiotemporal analysis demonstrates that most of the schistosomiasis clusters disappeared, or that their significance declined over time. In traditional hotspots, e.g. Ntaruka-Kinoni HFSA in Burera District, rates have declined consistently. Indeed, this area changed from the most important cluster during sub-period I into a secondary cluster in the following two sub-period. Over the course of three years, the selected areas (Figure 21) were covered by repeated MDA campaigns, guided by the district prevalence map of the nationwide prevalence mapping survey. The approaches initiated by the NTD control programme were also complemented by the great positive impact of improved socio-economic conditions (such as increased number of

school attendances, access to improved water for domestic use, proper sanitation, wearing shoes, etc.) (Leuchowius, 2014; MINECOFIN, 2015; United Nations, 2000).

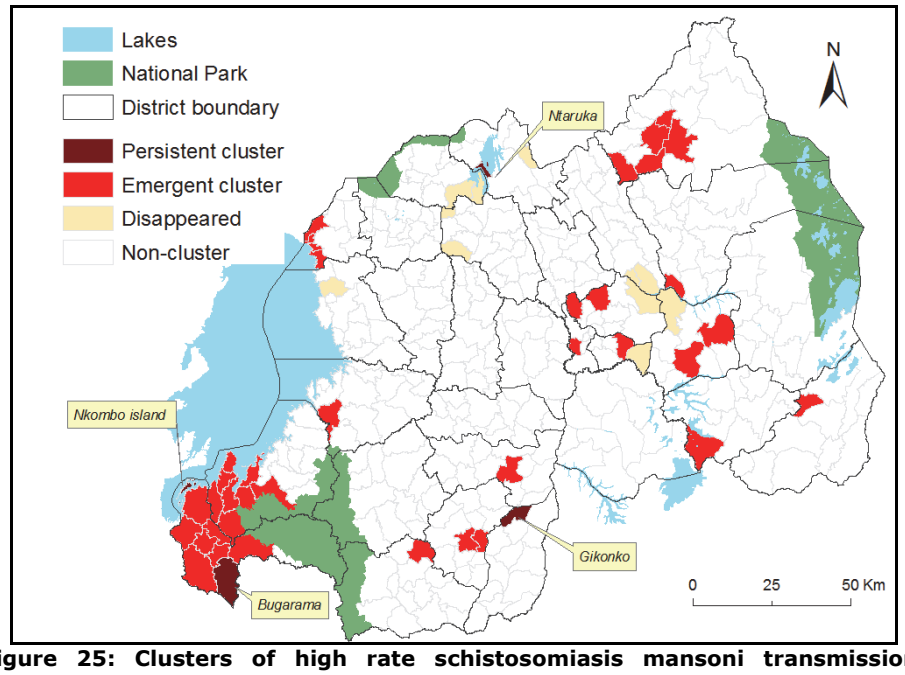

Figure 25: Clusters of high rate schistosomiasis mansoni transmission. Persistent clusters; emerging clusters appearing since 2008; clusters which disappeared after control measures had been implemented (sand color) and non-endemic areas.

Although the mapping exercise was generally of great help in guiding the NTD control initiatives, it did not capture all high-prevalence locations. The most probable reason for this was that spatial resolution of the district-level baseline map could have been better. The district-based approach was too aggregated, and therefore some high rate transmission areas were not detected. For example, Nkombo Island and neighboring areas are located in Districts classified as low endemic, while Ruberanziza et al. (2015) identified it as one of the most important foci in Rwanda. Therefore, the areas become the most important cluster in sub-period III. Since the districts Rusizi and Nyamasheke were falsely considered among the least endemic districts with prevalence rates of 0,7 % and 4.5 %, respectively (TRAC+, 2008), they did not benefit from control measures during the first three years of the interventions (Figure 21).

Despite, the substantial impact of the well-managed NTD intervention programme, our analysis revealed that the schistosomiasis rates remain high and that there are also new emerging clusters. Thus, the eradication of the

disease is difficult since the transmission factors are still there. The persistent and emergent clusters may be associated with various and changing risk factors as discussed by Steinmann et al. (2006) in a systematic review of dam projects referring to other African countries that are also endemic for schistosomiasis. In Rwanda, the most important risk factors might be related to the extended and new irrigated areas as a result of the agricultural transformation in Rwanda (Malesu et al., 2010). Furthermore, conversion of wetlands for rice cultivation has increased from 4000 ha to 13000 ha between 2000 and 210 (Nabahungu & Visser, 2013). Consequently, two of four persistent clusters (Bugarama and Gikonko HFSAs) are both located in areas of the oldest (since the 1980's) and also the largest rice cultivation schemes (Rwanda, 2010). According to unpublished statistics from Ministry of Agriculture, Rusizi District hosted 20% of the total rice crops in Rwanda by 2012 and Bugarama District counts more than 6,000 of about 17,000 households engaged in rice cultivation. Furthermore, rice cultivation in Rwanda involves the most vulnerable community groups, women, and children; more than 45% of rice farming workers are women (Nkurikiye, 2016). The number of schistosomiasis infections may double or triple for a relatively small number of peoples in temporal shelters/villages of workers concentrated around such water development projects. Adjacent to most wetlands there is relatively flat areas that are easily settled and which lack hygienic conditions, such as tap water and sewage systems. The expansion of irrigated rice cultivation (Figure 26) may explain the expanding clusters in the Eastern Province, as illustrated throughout by Figures 24 and 25. The Eastern Province totaled 41,5 % of rice cultivation areas in Rwanda by 2012. The expansion of rice cultivation in the Eastern Province took place specifically within Nyagatare District, where rice cropping in wetlands along Muvumba River increased from 180 to 1,350 ha between 2000 and 2012. Also, the Eastern Province, a relatively semi-arid region of the country, a lot of dams have been constructed, and hillside irrigation projects initiated (Malesu et al., 2010); MINAGRI, 2012).

Other persistent clusters (Ntaruka-Kinoni and Nkombo HFSAs) are situated in close proximity of Kivu Lake and the adjacent lakes of Burera and Ruhondo. Most of the inhabitants in these areas are in permanent contact with the lakes when fetching water for domestic use, swimming, and fishing. This finding is in line with other studies highlighting the link between schistosomiasis and water contact (Tukahebwa et al., 2013; Mazigo et al., 2014). Likewise, most of the emerging endemic areas are in proximity to water bodies and man-made agricultural schemes, which supports the intermediate snail host and transmission (Adenowo et al., 2015).

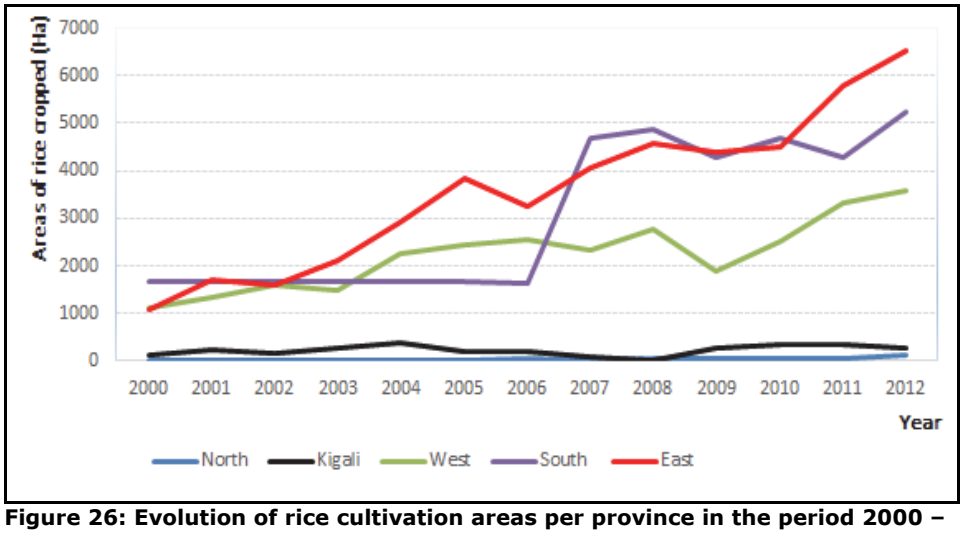

Figure 26: Evolution of rice cultivation areas per province in the period 2000 – 2012

Although the spatially structured data recorded at primary health facilities inform us about the distribution of disease transmission, the people tested constitute just the portion representing the seriously infected persons; the total number of infections is surely much greater. We advocate more actively monitoring for all newly emerging clusters using existing routinely collected data, to guide new intervention programs. Disease-conscious planning of land use in the future could also help to prevent further increase of the disease.

Although the level of aggregation of the data investigated is relatively small (at the HFSA level), they do not represent the exact geographical location of the infected patients for most HFSAs as unoccupied areas are not excluded (i.e.: protected areas). Thus, the coordinates linked to the HFSA centroid used may not reflect the reality needed/tracked by the spatial scan window applied. However, earlier comparative studies by Torabi and Rosychuk (2011) and Rashidi et al. (2015) show only minor differences between circular and flexible scan statistic methods in cluster analysis.

5.5 Conclusion

The analysis presented here has demonstrated that the presence of schistosomiasis in Rwanda varies across space and time. The impact of the NTD national control programme was clearly reflected by localized clusters that were either persistent, emerging or disappearing over time. The intervention initiative pushed hotspots of schistosomiasis transmission to the periphery but was not able to eradicate the disease. Persistent clusters with high schistosomiasis rates still occur, while emerging clusters show a linkage with new, large-scale agricultural schemes involving irrigation and rice

cropping. The results generated should be highly valuable for a spatially structured evaluation of risk factors associated with schistosomiasis and provide vital information for decision-makers and health planners. The methodology presented and outcomes generated contribute to a better understanding of the dynamics of schistosomiasis and are recommended for future prevention, control, and elimination interventions.

Chapter 6: Modelling schistosomiasis risk areas dynamics over time in Rwanda using zero-inflated Poisson regression

Abstract

The full count of infections can be hardly achieved. As an alternative, the recorded schistosomiasis cases has some areas with zero confirmed cases which do not imply that there are no infected persons. Furthermore, the standard statistical analysis, using exploratory or confirmatory spatial regression, are failing to account for the possibility of missing data in the records. Therefore, a spatiotemporal schistosomiasis risk modelling accounting for the imperfect case detection with varying environmental risk factors was the aim of this study. The Bayesian Hierarchical spatiotemporal model was applied to assess the schistosomiasis risk areas dynamics over time and to develop future scenario by 2050. The model used incidences data recorded at primary health facility level and environmental data. The current trend of identified risk factors was used to project their distribution up to 2050 for forecasting the disease risk pattern. The zero-inflated Poisson model shows an explosive increase of relative risk of the schistosomiasis over one decade. Furthermore, the change between relative risk of 2009 and forecasted risk by 2050 computed persisting and emerging areas with high relative risk of schistosomiasis. The risk of schistosomiasis transmission is 0.69, 0.29 and 0.50 time (to say 69%, 29%, and 50%) higher into areas with rice cultivation, proximity to rice farms, and proximity to a water body respectively. The prediction and forecasting maps provide a valuable tool for monitoring schistosomiasis risk in Rwanda and planning future disease control initiatives.

Keywords: Schistosomiasis, Zero-inflated Poisson model, environmental factors, infection cases, Spatial

6.1 Introduction

Schistosomiasis, an acute and chronic disease, is a well-known environmental-related disease (Hu et al., 2016). People are infected during routine agricultural, domestic, occupational and recreational activities which expose them to infested water (WHO, 2010b). Given the increasing number of people living in rapidly extending irrigation schemes or in close proximity to large dam reservoirs, a large number of the community are at high risk of schistosomiasis (Steinmann et al., 2006). A better understanding of where and when schistosomiasis transmission is likely to occur would enable more effective monitoring strategy and control measures.

In Rwanda, schistosomiasis mansoni or intestinal schistosomiasis has an overall country prevalence of 2.7%, per district, ranging from 0 to as much as 69.5 % among school children, constitutes a significant public health problem (TRAC+, 2008). Therefore, the Neglected Tropical Disease (NTD) control program, was established in 2007 with a special emphasis on schistosomiasis mansoni. The schistosomiasis control initiative has developed capacity in disease diagnosis and treatment complemented by MDA campaign within significantly endemic areas. Furthermore, sentinel sites (12 schools and 2 villages) were selected based on nationwide prevalence mapping, for schistosomiasis surveillance (Ruxin & Negin, 2012). However, high rates of infection in traditionally endemic areas are persisting (Ruberanziza et al., 2015) and new schistosomiasis foci within previously non-endemic zones are appearing (Isabwe et al., 2012; Ruberanziza et al., 2010).

Therefore, a detailed spatiotemporal disease risk assessment can strengthen the control strategies and interventions for S. mansoni. To do this, we have investigated the spatial distribution of S. mansoni using spatially structured and systematically recorded incidence data (chapter 3). The data was found to be valuable in detecting a spatial pattern of this highly focalized disease with dynamic distribution in time and space such as schistosomiasis (chapter 5), similarly to others water based diseases (Odoi et al., 2003; Osei & Duker, 2008; Vega-Corredor & Opadeyi, 2014). The study demonstrated that schistosomiasis was not distributed randomly and is significantly associated with wetlands and their use (agricultural). We further utilized spatial Scan statistics (Kulldorff, 1997) to detect spatiotemporal clustering of schistosomiasis during 2001 – 2012 (Nyandwi, Veldkamp, Osei, & Amer, 2017). This analysis showed a dynamic distribution of schistosomiasis in space and time.

However, the standard statistical analysis, using exploratory or confirmatory spatial regression, are failing to account for the possibility of missing incidence data in the records or to include risks for those areas that have

zero cases (Law & Quick, 2013). Areas with zero confirmed cases do not imply that there are no infected persons. Thus, the interpretation of results may be unclear since associated factors can be related either to the occurrence of the disease or to the (in) effectiveness of the surveillance data. Prior investigations demonstrated that schistosomiasis increase can be the result of the construction of dams or irrigation schemes (Yapi et al., 2005). In Rwanda, many small and large scale wetland resources development projects are rapidly taking place (Malesu et al., 2010). Also, climatic factors (warm and humid conditions) have their influence on disease transmission (Codjoe & Larbi, 2016; McCreesh & Booth, 2014), but they also have a significant impact on Rwandan wetlands distribution (Nyandwi et al., 2016). Therefore, focalized schistosomiasis risk has not yet predicted or forecasted at a detailed spatial scale in Rwanda. Schistosomiasis risk modeling accounting for the imperfect case detection with varying environmental risk factors in space and time is urgently needed.

Hence, modeling the spatiotemporal risk areas of schistosomiasis with consideration of wetland ecosystem and climatic factors dynamics can play an important role in the monitoring of schistosomiasis in Rwanda. Such a model should go beyond the spatial and temporal dimensions separately and consider the combined space-time interaction dimensions. This study investigates the dynamic spatial pattern of schistosomiasis mansoni risk in Rwanda using a Bayesian model counting for false zero cases and (ii) to develop a scenario of future risk areas by the year 2050 using policy based assumptions. The model could provide an efficient tool for specific interventions in the future and should be adopted to others human diseases with recorded routine health data at detailed spatial scale.

6.2 Materials and Methods

Study area

Rwanda, a small, densely populated, landlocked country of 26,338 km2 in the Great Lakes Region of central-eastern Africa, is administratively divided into five provinces, 30 districts and 416 sectors (MINALOC, 2005) and has a population of about 11. 5million (2016). The majority of the population are employed in agriculture (85). Rwanda has a relatively dense and dynamic hydrological network with many lakes and rivers and numerous floodplains and wetlands covering 10% of the land surface (Nyandwi et al., 2016). Thus, the wetland reclamation and irrigation project for agriculture transformation and modernization are pillars strategies to achieve the MDG 1 – on eradicating extreme poverty and hunger.

Parasitological data description

The confirmed cases of schistosomiasis mansoni recorded per primary health facility (covering more or less one administrative sector) were provided by the Malaria and Other Parasitic Diseases Division of the Rwanda Biomedical Centre (RBC/M&OPD). The dataset consists of laboratory confirmed cases from all public and faith-based health facilities for the period January 2001 to December 2009. The reported number of confirmed cases of schistosomiasis by each primary health facility during 2001 – 2009 is likely to be structurally zero-inflated. A health facility service areas (HFSA) where no cases were detected could indeed be an HFSA where no infection occurred (true zero) or it could be an HFSA where at least one infection occurred but none were reported or diagnosed (false zero). Therefore, methods that can deal with data insufficiency are needed (Gelman, 2008). Not accounting for such insufficiency can lead to unstable estimations of infection patterns. The spatial zero-inflated Poisson methods are becoming a common tool for disease risk modeling (Vergne et al., 2014) and forecasting (Hu et al., 2016).

Statistical analysis and model inference

Zero-inflated Poisson model

Let Y_{it} be a random variable of S. mansoni outcomes with realizations y_{it} for $i = 1,...,367$ HFSAs for $t = 1,...,9$ years. It is assumed that the data are generated by a Poisson distribution. Initial exploratory analyses indicate the presence of numerous zeros with $\Pr(y = 0) = 0.29$ (See Figure 27). The number of zeros and the heterogeneous distribution of the positive counts imposes competing influences when the standard Poisson model is applied. Alternatively, we use zero-inflated Poisson likelihood to fit a three stage Bayesian Hierarchical spatiotemporal model for *S. mansoni* risk. In this context, we model y_{it} as a mixture of Poisson distribution and a point mass at 0. Thus, when there are no cases, i.e. $y_{it} = 0$, we assume such zero counts are generated from a Poisson distribution with probability $1-\Phi_i$, and as ***sampling zeros*** with probability Φ_i. Formally, this can be expressed as

$$\Pr\left(Y_{it} = y \middle| \Phi_i, \lambda_i\right) = \begin{cases} \Phi_{it} + \left(1-\Phi_{it}\right)e^{-r_{it}p_{it}} & \text{if } y = 0 \\ \left(1-\Phi_{it}\right)\dfrac{e^{-r_{it}p_{it}}\left(r_{it}p_{it}\right)}{y!} & \text{if } y = 0,1,2,... \end{cases} \quad (1) \textbf{\textit{(Eq.10)}}$$

where r_i and p_i are the risk and population at the location i, respectively Here, the zeros are a mixture of two distributions, the binary and the Poisson

distribution that includes zeros. This distribution expresses the idea that an observation can be zero even though the disease is present. This is appropriate when there are detectability/diagnostic problems and/or deficiencies in surveillance systems, and the reluctance of infected individual to seek medical attention. We specified an alternative model assuming that only structural zeros are present and interpret the zeros as arising from only the binary distribution. In that case, we have the expression

$$\Pr\left(Y_{it}=y\middle|\Phi_i,\lambda_i\right)=\begin{cases}\Phi_{it}+\left(1-\Phi_{it}\right)e^{-r_{it}p_{it}} & \text{if } y=0\\ \left(1-\Phi_{it}\right)\dfrac{e^{-r_{it}p_{it}}\left(r_{it}p_{it}\right)}{y!} & \text{if } y=1,2,\ldots\end{cases} \quad (2)$$ ***(Eq.11)***

At the process stage, our interest is to model the risk of infection r_{it} as a latent random field conditional on y_i being either sampling or structural zeros. Thus

$$\ln r_{it}=\beta_0+\Sigma f_p\left(x_{ip}\right)+\Sigma z'_{itk}\gamma_k+u_i^{CAR}+\upsilon_i^{iid}+t_t^{RW1}+\zeta_{it} \quad (3)$$ ***(Eq.12)***

where β_0 is the intercept, f_p are appropriate smooth functions of $p=1,..,P$ continuous covariates x_{ip}, and z'_{itk} is a vector of $k=1,..,K$ categorical covariates with associated parameters γ_k. We specified $u_i^{CAR}\sim\text{ICAR}\left(w,\sigma_u^2\right)$ as intrinsic conditional autoregressive (ICAR) process with variance σ_u^2 and the random intercepts $\upsilon_i^{iid}\sim N\left(0,\sigma_\upsilon^2\right)$ as zero-mean Gaussian process with variance σ_υ^2, Here the weight matrix *w* represent the spatial neighborhood structure. By convention, and as applied in most studies, two HFSAs are assumed neighbors if they share a common boundary. Thus we defined neighborhoods as adjacent HFSAs with simple binary adjacency weights, i.e., $w_{ij}=1$ if areas i and j share a common boundary and $w_{ij}=0$ otherwise. For the continuous covariate x_{ip} of *M* equally spaced knots $x_p^1<x_p^2<\cdots<x_p^M$, we specified the nonlinear function $f_p\left(x\right)=\xi_m$, $m=1,\ldots M$, with second-order random walk priors $\xi_{pm}\sim N\left(2\xi_{p,m-1}-\xi_{p,m-2},\sigma_\xi^2\right)$ and non-informative priors $p\left(\xi_1\right)\propto 1$ and $p\left(\xi_2\right)\propto 1$. Here variance parameters σ_p^2 control the amount of smoothing. Also for the temporal trend, we write $f\left(t\right)=\rho_m$ and specify second-order

random walk prior $\rho_m \sim N\left(2\rho_{m-1} - \rho_{m-2}, \sigma_\rho^2\right)$ and non-informative priors $p\left(\rho_1\right) \propto 1$ and $p\left(\rho_2\right) \propto 1$. Lastly, we specified $\zeta_{it} \sim N\left(0, \sigma_\zeta^2\right)$ as zero-mean random space-time interaction effects. Let $\boldsymbol{\xi} = \left\{\xi_{pm}\right\}$ be a vector of PM parameters, $\boldsymbol{\rho} = \left\{\rho_q\right\}$ be a vector *Q* temporal trend parameters, $\boldsymbol{\gamma} = \left\{\gamma_k\right\}$ be a vector of *K* categorical parameters, and $\boldsymbol{\zeta} = \left(\zeta_{it}\right)$ be a vector of 3303 space-time parameters. The full Gaussian latent field is then $\boldsymbol{\psi_1} = \left\{\beta_0, \boldsymbol{\xi}, \boldsymbol{\gamma}, \boldsymbol{\rho}, \mathbf{u}, \boldsymbol{\upsilon}, \boldsymbol{\zeta}\right\}$.

At the third stage of the Bayesian hierarchical model, we treat the precision/variance and the fixed parameters as unknown and assign prior distribution for their joint density. We specified non-informative priors for the random walk parameters $p\left(\xi_1\right) \propto 1$, $p\left(\xi_2\right) \propto 1$, $p\left(\rho_1\right) \propto 1$ and $p\left(\rho_2\right) \propto 1$. For β_0 and parameters in the vector, $\boldsymbol{\gamma}$ we assumed non-informative Gaussian distribution with zero mean and precision 10^{-5}. Thus $\beta_0 \sim N\left(0, 10^{-5}\right), \boldsymbol{\gamma} \sim N\left(0, 10^{-5}\right)$.

For the precision parameter,

$\tau_j = 1/\sigma_j^2$ we assumed $\tau_j \sim \log Gamma\left(1, 0.00005\right)$, $j = \xi, u, \upsilon, \rho, \zeta$ as prior distributions. Following the Bayesian paradigm, we aim to determine the posterior distribution of the unknown parameters based on their prior distribution. Let $\boldsymbol{\psi}_2 = \left\{\sigma_\xi^2, \sigma_u^2, \sigma_\upsilon^2, \sigma_\rho^2, \sigma_\zeta^2\right\}$ be a vector of all variance unknown parameters, one can simulate samples from the posterior density $p\left(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \middle| \mathbf{y}\right) \propto p\left(\mathbf{y} \middle| \boldsymbol{\psi}_1, \boldsymbol{\psi}_2\right) \times p\left(\boldsymbol{\psi}_1, \middle| \boldsymbol{\psi}_2\right) \times p\left(\boldsymbol{\psi}_2\right)$ using Markov Chain Monte Carlo (MCMC) simulation. However, in this study, we generated the samples from the posterior distribution using the Integrated Nested Laplace Approximation (INLA) (Rue and Martino, 2007). INLA is an emerging alternative to the MCMC that provides fast and accurate estimates of the posterior marginals through Laplace approximation, a deterministic algorithm proposed by Rue and Martino (2007). Details about this approach and its applications can be found elsewhere (Blangiardo and Cameletti, 2015; Rue and Martino, 2007; Rue and Held, 2005).

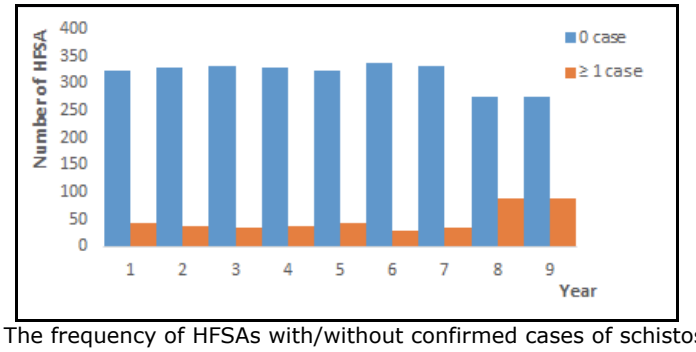

Figure 27: The frequency of HFSAs with/without confirmed cases of schistosomiasis.

Associated environmental variables

The variables used for modeling were extracted per HSFA units. Rice cultivation (*Rice*), proximity to rice farms (d_{Rice}), and proximity to water (d_{Water}) bodies as categorical variables, whiles wetland proportion (Swetnam & Reyers), rainfall (*Rain*), and temperature (*Temp*) were modeled as continuous variables. The variable *Rice* is an indicator for rice cultivation in an HFSA. The variables d_{Rice} and d_{Water} are HFSA within a 5km radius of rice farms and water bodies, respectively. The cut-off value of 5 km ascends from studies on schistosomiasis (Handzel et al., 2003; Kabatereine, Brooker, Tukahebwa, Kazibwe, & Onapa, 2004). Detail descriptions of how these variables were extracted can be found in our previous study (Nyandwi et al., 2016). The combined categorical and continuous variables were mapped as illustrated in Figure 28.

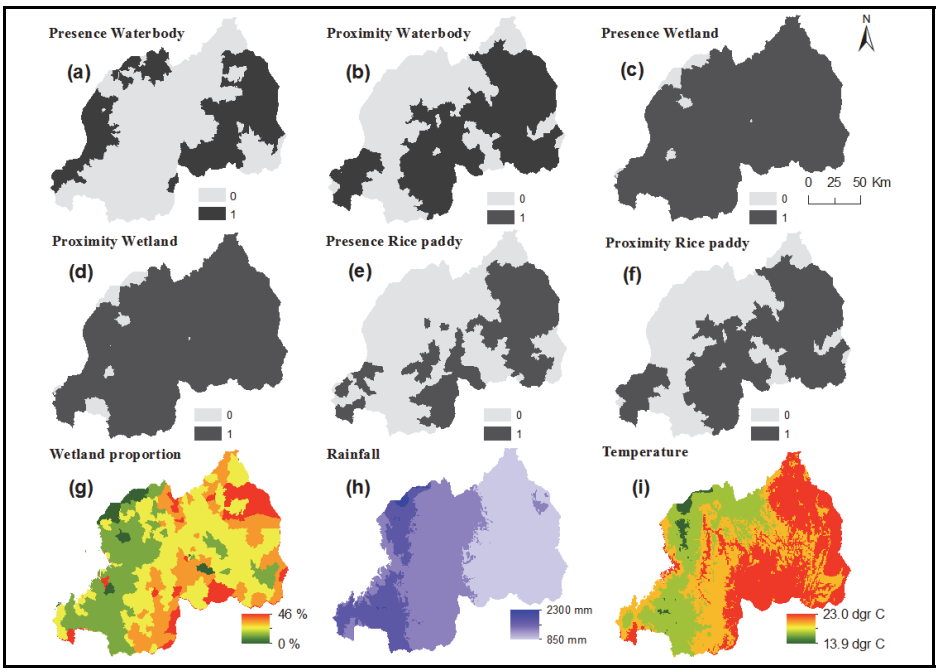

Figure 28: Risk factors maps. Categorical (from **a** to **f**) and continuous (from **g** to **i**) covariates.

Model implementation and performance evaluation

We implemented two models, model 1 and model 2 in the R-INLA package and discussed their statistical and substantive grounds. We specified the model with equation 10 as model 1 and the model with equation 11 as model 2. The full model was expressed as

$$\begin{aligned}\ln r_{it} = \beta_0 + f_{Wet}(Wet) + f_{Rain}(Rain) + f_{Temp}(Temp) \\ + Rice' \cdot \gamma_{Rice} + d'_{Rice} \cdot \gamma_{d_{Rice}} + d'_{Water} \cdot \gamma_{d_{Water}} \\ + u_i^{CAR} + \upsilon_i^{iid} + t_t^{RW1} + \zeta_{it}\end{aligned}$$

We compared the predictive performances of models 1 and 2 using the deviance information criterion (DIC) suggested by Spiegelhalter et al (2002). The $DIC = \bar{D}(\theta) + p_D$ is a two-term composite parameter of which the posterior mean deviance $\bar{D}(\theta)$ measures the fit to the data, whiles the effective number of parameters p_D measures the model complexity. In this regard, the smaller the DIC value the better the model fit and predictive performance.

Also, the relationships between observed and predicted relative risks were compared. A quasi-validation was also done by plotting observed and predicted relative risks during the study period (2001 – 2009).

Projection of future spatial trend of best-predicting risk factors

An important objective of the study was also to forecast schistosomiasis risk based on both models. Should the current trends of the risk factors continue up to 2050, what would be the spatial distribution of the risk of schistosomiasis? In order to answer this, the extra dataset was generated for the risk factors in order to produce the future scenario of schistosomiasis risk areas. (i) Rice paddy areas: The extent of wetlands to be reclaimed for rice paddy was defined by considering the national need in rice commodity and current yield, the rice cropped areas should increase from 12000 ha in 2010 to 55 000 ha (MINAGRI, 2013) and that may be achieved by 2050 considering the average increase of 1100 ha per year during last fifteen years. The spatial location of the defined rice paddy areas was delineated using the irrigation master plan (MINAGRI & ICRAF, 2010) and agriculture statistics and GIS data (unpublished). (ii) Wetlands extent: Using the probability map of Rwandan wetlands (Nyandwi et al., 2016), the extent of wetland by 2050 was projected. Climatic factors: The projection of future climate for Rwanda indicate a trend towards a warmer and wetter climate and mean temperature for 2050 was projected using a rate of 0.35oC increase per decade (McSweeney, 2011; Muhire & Ahmed, 2016). The rainfall records show a very small trend of an increase of 4mm rainfall amount (Muhire & Ahmed, 2015).

6.3 Results

Figure 29 displays the modeled endemic areas of schistosomiasis corresponding to HFSAs in Rwanda. The annual relative risk (RR) of schistosomiasis depicts the dynamic changes during the period of study (2001 to 2009). The spatial pattern of relative risk is quite dynamic across the country during the study period. But, in general, 340 HFSA have had less than 1 case of schistosomiasis per 10,000 people for most of the years. While significant risk (i.e., RR> 1) was predicted for about 37 HFSA. Very significant risk (RR > 3) are observed mostly within 3 HFSA.

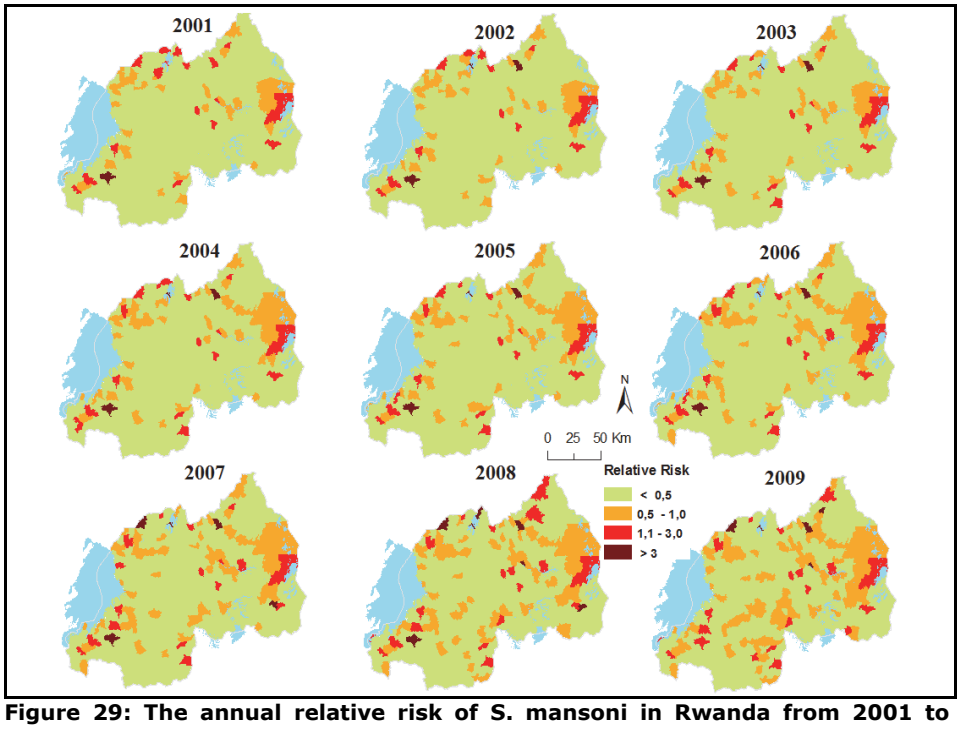

Figure 29: The annual relative risk of S. mansoni in Rwanda from 2001 to 2009.

As summarized in Table 16 and illustrated in Figure 30, schistosomiasis risk is positively correlated with the presence of rice cropping, proximity to rice cropping areas, water body, wetland cover proportion, and rainfall. Furthermore, is the disease negatively correlated with temperature. The identified variables are associated either with the predicted occurrence of the disease or with the number of confirmed cases (given disease occurrence) reported for each HFSA.

Table 16. Posterior estimates of model parameters with ZIP distribution of two models of Schistosomiasis risk in Rwanda, 2001 to 2009

	Model 1			**Model 2**		
Parameters	**Mean**	**$Q_{0.025}$**	**$Q_{0.975}$**	**Mean**	**$Q_{0.025}$**	**$Q_{0.975}$**
β_0	0.000134	0.000069	0.000263	0.0000785	0.00003	0.00019
γ_{Rice}	1.397	0.882	2.2071	1.688	0.989	2.891
$\gamma_{d_{Rice}}$	1.289	0.765	2.166	1.292	0.669	2.485
$\gamma_{d_{Water}}$	1.322	0.828	2.098	1.498	0.837	2.693
DIC		4970.10			4547.91	
p_D		296.45			336.84	

The model two performed best using the deviance information criterion (DIC). This model assumes that reported zeros cases for some HFSA were false zeros.

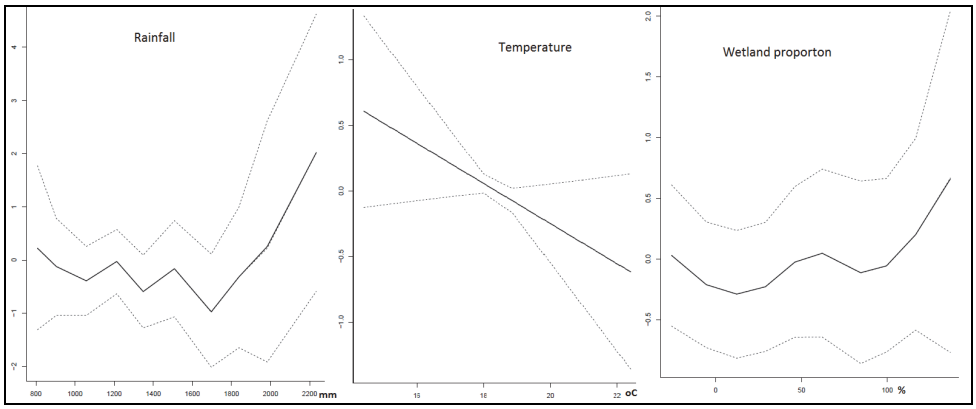

Figure 30: Model outputs from continuous variables

In Figure 31, a strong relationship between observed and modeled RR can be observed, with a coefficient of determination of 0.9. The relationships increase the confidence in formulating the assumptions around the model outputs for forecasting the RR by 2050.

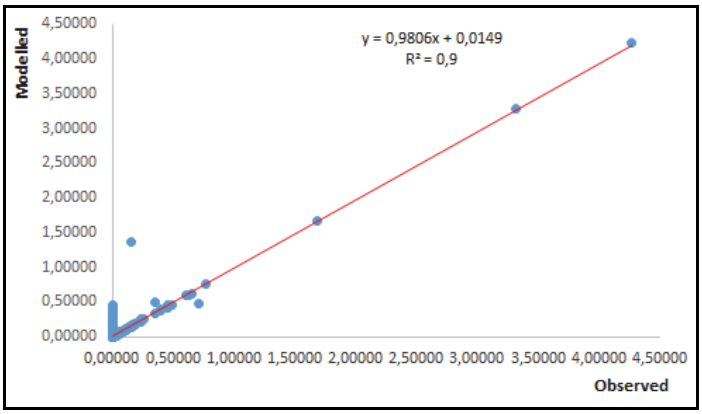

Figure 31: Scatter plot and trend line of Observed against modeled RR of schistosomiasis between 2001-2009

Figure 32 illustrate the probable spatial pattern of schistosomiasis considering the current temporal trend of the disease risk and projected trend of changes of associated risk factors by 2050. The endemic HFSA (*i.e.*, RR> 1) will shift from 26 in 2009 to 46 HFSA by 2050. While the very significant risk of schistosomiasis (RR > 3) will extend to nine HFSA by 2050 from five in 2009. In general, by 2050 the relative risk will only decrease within three HFSA,

while there is a general increase in risk ranging from 5 to 98% as illustrated by change map of schistosomiasis risk (Figure 32b).

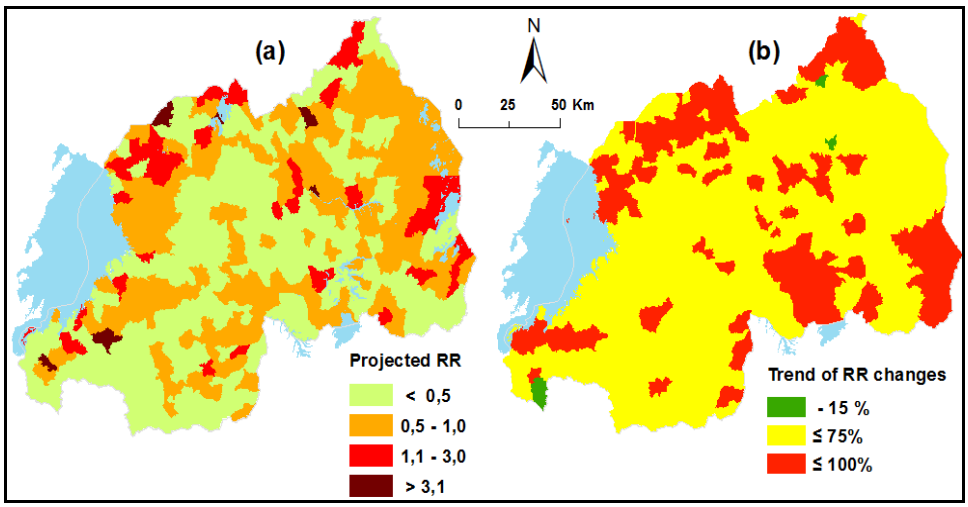

Figure 32: The future relative risk of schistosomiasis. Projected relative risk by 2050 (**a**) and change detected between 2009 and 2050 (**b**).

6.4 Discussion

The zero-inflated Poisson regression model shows an explosive increase of relative risk of the schistosomiasis over one decade. Furthermore, the change between relative risk of 2009 and forecasted risk of 2050 computed an increasing trend. The number of areas with significant risks will double by 2050 as illustrated by disease risk areas change the map in Figure 32b. Although a lot have been achieved by schistosomiasis control initiatives since 2008 (Ruxin & Negin, 2012), schistosomiasis risk areas are still unsteady in space and time. That situation confirms the high level of uncertainty for modeling schistosomiasis risk pattern with linear regression approach. We also compared the two models and DIC results indicated that model 2, considering the existence of false zeros in disease cases, was the best fitting model. That was in line with others studies considering the Bayesian model to be efficient as a nonlinear model, which can even detect the linear relationships if it does exist, for an imperfect disease case detection situation (Hu et al., 2016; Vergne et al., 2014).

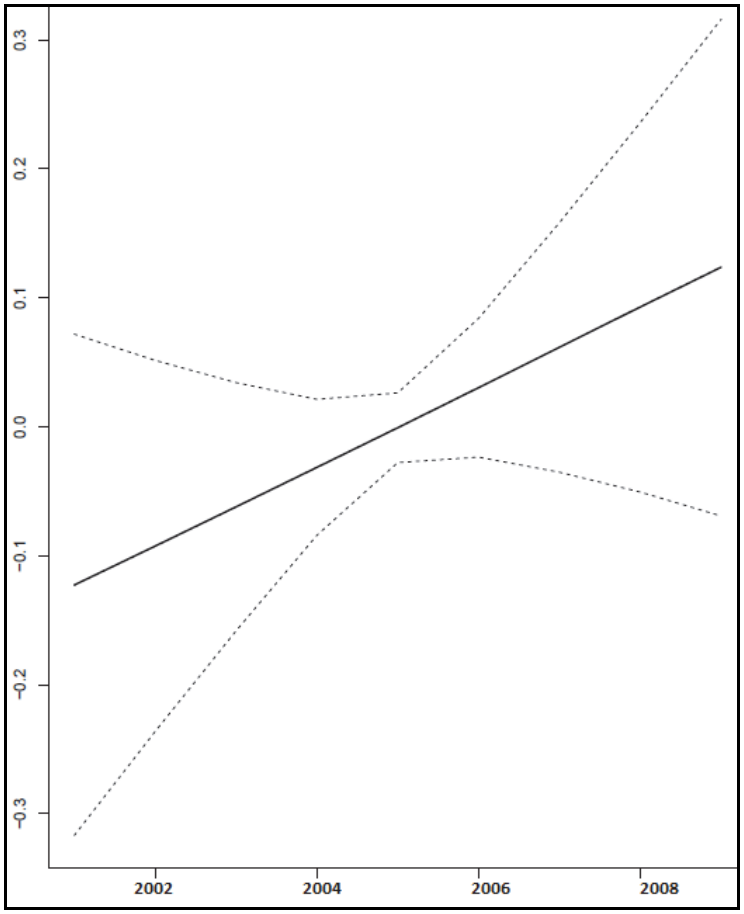

Figure 33: An increasing distribution of schistosomiasis cases over time

The increasing spatial variation of schistosomiasis case can be explained by statistically significant association with water and rice related environmental factors, as summarized by mean values in Table 16. The risk of schistosomiasis transmission is 0.69, 0.29 and 0.50 time (to say 69%, 29%, and 50% higher into an HFSA with rice cultivation (Rice), proximity to rice farms (d_{Rice}), and proximity to a water body (d_{Water}) respectively. The connection implies that the schistosomiasis infection will tend to be higher where opportunities to get in contact with contaminated water are plentiful (Steinmann et al., 2006). Infected people by Schistosomiasis mansoni are predominantly rural population whose water contact is linked to working activities, agriculture (rice crops), fishing (even in a boat), watering of cattle, trading, crossing, to domestic activities (water collection, washing of utensils and clothes), to personal activities (bathing), or to recreational activities (playing, swimming) (Handzel et al., 2003; Mote et al., 2005; Mupfasoni et al., 2009). Furthermore, the climatic factors add to the risk factors because

warm and humid conditions stimulate several soils and water-associated diseases (Degallier et al., 2010). Thus, different direction of correlation between schistosomiasis risk temperature is in line with the fact that extremely warm conditions increase vividly the evapotranspiration and risk for dry conditions halting the susceptibility of snails to infection with schistosomes (Knight et al., 2015).

The detected trend for disease pattern with significantly associated non-fixed environmental factors was also maintained in forecasting the disease by 2050. Climate observations over the past 40 years (McSweeney, 2011) allowed projection for the coming 40 years (2009 – 2050) and their significant correlation with observed disease risk might have to be maintained in disease risk forecast. Likewise, the projection of a water body, wetlands, and rice cropping related factors was based on local realities. However, our results should be used with caution because it was assumed that others potential risk factors remained constant over time. The study period was subject to an important NTD control program initiated in 2007 and the impact of improved disease diagnostic and drug treatment was reflected in schistosomiasis spatiotemporal distribution after effective implementation of the program (Nyandwi et al., 2017). But the same study has shown that schistosomiasis mansoni is rebounding and persisting in its traditional endemic areas and emerging in new areas. Surprisingly, the traditionally endemic area in extreme southwest (Bugarama) will be among three HFSA with drastically decreased risk by 2050 (Figure 32b). The value in time and space of forecasted risk should be used with some caution by public health professional and decision-makers in planning for disease control interventions (Clements, Firth, et al., 2010).

6.5 Conclusion

We selected a Bayesian statistical model adequate for a dynamic spatiotemporal distribution of schistosomiasis risk using routine health data of disease cases records with areas having zeros either by chance correct or as false zeros. Structured additive models combining continuous and categorical variables revealed increasing spatiotemporal patterns of schistosomiasis risk over time, significantly associated with presence or proximity to a water body and rice paddy, wetland cover proportion and climatic variables. Furthermore, the forecasting results exposed a persisting and emerging HFSA with a high relative risk of the disease. The prediction and forecasting maps provide a valuable tool for monitoring schistosomiasis risk in Rwanda and planning future disease control initiatives.

Chapter 7: Synthesis

In the introduction of this thesis, it was stated that Rwandan wetlands are ecosystems of critical importance to almost 12 million people, generally engaged in smallholder self-sufficiency agriculture. In recent years, wetlands are increasingly used for (intensive) agricultural production to ensure food security of the growing population. A proper appreciation of the potential economic benefits, and sustainable use, of wetlands requires an accurate delineation and typology of wetlands. Such a demarcation and characterization was absent at the start of this research. Also unknown was if and how intensified agricultural use of wetlands (a benefit) could have a negative impact (a cost) on the health status of communities that depend on their produce (Lichtenberg, 2002). This thesis aimed to assess the association between the wetland environment on the one hand, and schistosomiasis and Soil Transmitted Helminths (STH) incidence on the other. The starting point was to model and map wetlands and their characteristics both at a national and regional scale. After this, potential environmental and socio-economic risk factors were identified for STH and schistosomiasis transmission at a detailed spatial scale. Analyses were based upon routinely collected data of confirmed cases recorded at primary health facilities. Confirmed case data were transposed to health facility service areas (HFSA) which could accurately be demarcated and formed the basic spatial unit of analysis.

7.1 Summary of research findings

As discussed in detail in Chapter 2, the interactions between geomorphologic setting (elevation, slope), hydrological conditions (contributing area) and climatic factors (temperature, rainfall) are fundamental for characterizing Rwandan wetlands. Use of these factors in logistic regression analysis generated wetland spatial distribution probability maps with a calibration accuracy of 87,8% at national level, and an even higher accuracy of 98% at the subnational level. The national and sub-national models (10 agro-ecological zones and 10 water catchments) were used in a sensitivity scenario analysis to assess future wetland location probability as a result of changes in temperature and rainfall. A 1% increase in temperature would cause a net probable wetland area decline of 12.4%. A 1% increase in rainfall would cause a net probable wetland increase of 1.6%. Our analysis shows that wetlands in the central part of Rwanda are potentially the most sensitive to climate change. Hence, future wetlands use planning should be based on a multi-scale approach, to generate location-specific wetland use plans.

The wetlands probability map was an essential input for the subsequent assessment of the spatial relationships between STHs and schistosomiasis incidence and risk factors associated with intensified agricultural use of wetlands. To overcome the low granularity mapping of the 2007-2008

nationwide cross-sectional prevalence survey, we evaluated the value of using available routine health data recorded at primary health facility level. First, a nationwide primary health facility location map, consisting of all public health facilities, was generated. After that 376 health facility service areas (HFSA) were spatially delineated using cost allocation spatial analyst techniques, extensive field work, and consultations with GIS and public health professionals from the various administrative districts and the Rwanda Biomedical Centre (RBC). We then evaluated the spatial pattern of STH and schistosomiasis cases at HFSA and District level and related them with potential risk factors using an empirical modelling approach (see Chapter 3 and Chapter 4). The comparison of prevalence and incidence data showed that they were strongly correlated for schistosomiasis (R^2 = 0.79) and ascariasis (R^2= 0.63). Subsequent analysis showed that the spatial pattern of disease incidence has a localised distribution and is therefore better monitored at HFSA then at the much larger District level. Also, a significant positive relation was identified between schistosomiasis incidence and presence of and proximity to rice cropped areas, rainfall and particular soils properties. For STH incidence, population density and the proportion of rural residents were found to be significantly correlated.

In Chapter 5 a spatio-temporal analysis was conducted to identify variations in clustering of schistosomiasis incidence at HFSA level across space and time using SaTScan statistics. Our study revealed an increasing temporal trend of schistosomiasis, with persisting, emerging and disappearing clusters in time and space. The spatiotemporal pattern allowed us to distinguish three sub-periods which can be linked to disease control program initiatives. A seven year sub-period with a stable number of cases (2001 – 2007), was followed by an abrupt increase of recorded cases during the three of control program implementation phase. This was followed by declining and disappearing clusters when the program was in full place. Our findings illustrate that the spatial and temporal distribution of schistosomiasis incidence is very dynamic which makes elimination difficult to achieve. Identified spatio-temporal dynamics should, therefore, be considered in future disease control activities.

Given that the spatiotemporal analysis revealed many changes with many HFSAs with zeros counts, an empirical regression alone was considered insufficient. Modeling current and future risk areas of schistosomiasis in Rwanda was still needed before concluding this research. To achieve this, zero-inflated Poisson modeling was conducted in Chapter 6. The zero-inflated Poisson model that was generated shows an explosive increase of relative risk of schistosomiasis transmission in one decade. Furthermore, changes between the relative risk of 2009 and forecasted risk in 2050 indicated persisting and emerging areas with high relative risk of schistosomiasis transmission. Model outcomes indicate that the risk of schistosomiasis

transmission is 69% higher in areas with rice cultivation, 29% higher in areas in close proximity to rice farming, and 50% higher in areas close to water bodies.

7.2 Reflection on research findings

This section reflects on the main study findings. The reflection is structured around five subjects: (1) spatial scale sensitivity of wetland and disease pattern modelling, (2) climate sensitivity of wetland probability maps and disease distribution, (3) value of routinely collected health data, (4) spatiotemporal dynamics of wetland reclamation (i.e. rice cropping) and schistosomiasis incidence, and (5) recommendations for future wetland management and disease monitoring planning.

1. Modelling wetlands and disease patterns is sensitive to spatial scale

Spatial scale

This thesis has evidenced that wetland development and human health impact models and analysis outcomes are scale sensitive. Thus, future (sustainable and integrated) wetlands management in Rwanda should be location specific. Probability modeling of wetlands in Rwanda using logistic regression at national and regional level confirmed the explanatory power of topographic characteristics and climatic factors. Topographic characteristics are relatively stable landscape features, whereas climatic factors are more dynamic and have varying influence at different scales of analysis and in different zones of the country. Wetland occurrence prediction reached accuracy levels ranging from 80% (within Akanyaru watershed) to 98% (within Buberuka Highlands AEZ and Mulindi watershed) as summarized in Table 3a &b in Chapter 2. Our findings were in line with other studies recognizing the changing impacts of drivers at different spatial scale (de Koning et al., 1998; Kok & Veldkamp, 2001).

The spatial pattern and spatial autocorrelation tests demonstrated the strong locality of schistosomiasis mansoni hotspots (Ruberanziza et al., 2015). Using data of confirmed cases from each health facility enabled a high-resolution statistical investigation of schistosomiasis hotspots. Prevalence survey data aggregated to District level result in an overestimation of areas of high transmission, as well as in non-identification of hotspots in Districts with overall low incidence rates. Using the routine health data also enabled identification of the spatial patterns and clustering of individual and combined STH infection (Chapter 4). Specific spatial patterns for each individual STH type were identified as follows: roundworm and whipworm were found to be localized with a significant spatial autocorrelation, while hookworm was

ubiquitous and randomly distributed geographically (Table 11). There are districts with high, moderate and low incidence rates of combined STHs cases. The study demonstrated that the STH cases are geographically clustered in areas smaller than District level which has been used to date to report STH occurrence

Temporal scale

The study also established that climate change (temperature and rainfall changes) will have spatially varying effects on wetlands occurrence.

The space-time analysis of schistosomiasis incidence patterns made that we could identify three periods related to the NTD control program. A period prior to establishment of the national NTD program characterized by relatively low levels and slightly varying numbers of cases. During the first three years of the NTD control program, the number of confirmed cases increased sharply. The increased number of cases can be explained by increased numbers of people seeking health care, as well as improved diagnostic skills and availability of laboratory facilities (Ruxin & Negin, 2012). At the same time new hotspots of schistosomiasis transmission, emerged in areas with rapid expansion of rice cropping areas. From 2010 onwards, the number of cases decreased due to increased awareness of local communities and strengthened public health interventions.

The interventions can now be directed geographically and over time much more accurately than the current prevalence approach can for dealing with spatiotemporal dynamic schistosomiasis clusters. Our spatial pattern modeling approach can even help in guiding an improved sampling strategy for prevalence surveys, which will be always needed for disease morbidity monitoring.

2. Wetlands, wetland use, and health risks are climate sensitive

Both wetland occurrence patterns and disease incidence patterns are climate sensitive in Rwanda. The research identified and verified the effect of climate factors on wetland patterns and on schistosomiasis and STHs patterns.

While the existing classification of Rwandan wetlands reflected the utilization and stages of their development (REMA, 2008), the interdependency of geomorphic setting, water source and its transport and hydrodynamics (Buis & Veldkamp, 2008a; Xie et al., 2011) were not considered. The climate change scenarios developed in Chapter 2 demonstrated that wetlands in particular zones in Rwanda are more sensitive to climate variations. The central zones of the country are the most sensitive to climate factors, as illustrated in Figure 34.

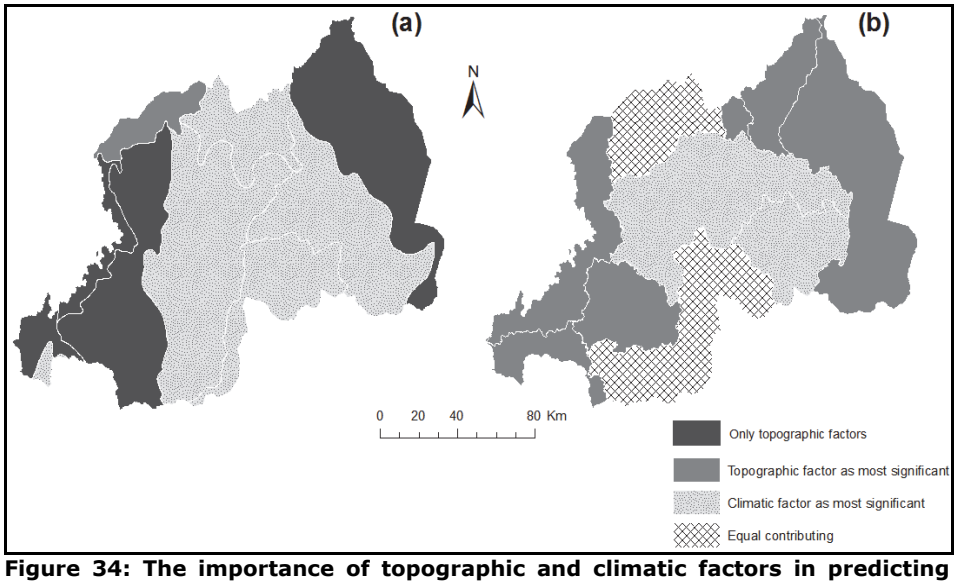

Figure 34: The importance of topographic and climatic factors in predicting wetland at (**a**) Agro-ecological zone level and at (**b**) Watershed level.

The multi-scale quantification of climate sensitivity, using different values for temperature and rainfall in the model illustrate that wetlands in Rwanda are highly sensitive to climate change. An increase of 1% in the average temperature can result in a reduction of more than 12% of current wetland coverage at the national level and can reach a maximum of 37% recession within Mayaga Plateau and Central Bugesera agro-ecological zones.

The localized influence of climatic factors on wetlands and human interaction with such locations, create new areas with good habitats for host snails. Thus, the consideration of climatic among others associated risk factors is of vital importance for assessing possible future human health impacts induced by wetland development projects. The empirical regression model confirmed a significant and positive correlation between schistosomiasis incidence and temperature and rainfall factors (Table 8 and Table 13). The advanced spatial statistical analysis using zero-inflated Poisson modeling, also confirmed that the predicted occurrence and the number of confirmed cases (actual occurrence) of *schistosomiasis mansoni* in Rwanda are positively correlated with temperature and rainfall among others risk factors. An additional step was made by forecasting the disease risk distribution by 2050 under the assumption of current policies and projected change trends of significantly correlated risk factors (as explained under the methodological approach of Chapter 6). The projection of future climate for Rwanda indicates a trend towards a warmer and wetter climate. Temperature is increasing with a rate of 0,35^{o}C per decade (McSweeney, 2011; Muhire & Ahmed, 2016). While rainfall records show an increase of 4mm (Muhire & Ahmed, 2015).

Consequently, disease eradication is not easily achievable since persistent clusters with high transmission rates still occur while emerging clusters show a relation with risk factors such as expanding irrigated agriculture and changing climatic factors.

3. Added value of routinely collected confirmed case data for identification and analysis of spatial patterns of schistosomiasis an STH

Usefulness of combining incidence and prevalence data for disease monitoring

Conventionally, NTD control programs, are guided by results of cross-sectional prevalence surveys aggregated to larger areas. The sample size taken much depends on available resources. As stated in the current mapping guidelines of the World Health Organization (Montresor et al., 1998), sampling sites are chosen on the basis of lottery methods. Prevalence rates obtained at each sample site are then extrapolated to larger areas (e.g. results from 3 to 4 surveyed school were aggregated to the district level during the nationwide NTD prevalence mapping of 2007/8). Therefore, morbidity control initiatives are relatively expensive and permanent monitoring of disease hotspots is problematic in resource constrained conditions.

This research revealed that routine health data can provide valuable information and the HFSA proved to be a suitable small-scale geographic unit for the analysis of schistosomiasis distribution. The study demonstrated a strong, location–specific, correlation between prevalence data and incidence rates of schistosomiasis (Chapter 3), with a strong coefficient of agreement as illustrated in Figure 11. Similarly, a high correlation was found for roundworm (R^2 = 0.63), one of the three STHs that represent a significant public health problem in Rwanda.

The six reasons justifying the high quality of recorded schistosomiasis incidence data from R-HMIS was explained in the methodology section of Chapter 3. The remaining question that needed to be addressed was to conduct a spatiotemporal modeling of schistosomiasis to identify current and future areas of high transmission risk in Rwanda. A standard linear regression analysis cannot take into account the existence of many false zeros of confirmed disease cases in many HFSAs and many changes of disease hotspots across space and time. Projecting future schistosomiasis risk could be done because of the good quality (spatially structured and systematically collected) of the routine health data, and small-scale data on irrigated areas, together with methods provided by current GIS and spatial statistics for epidemiological studies (Tsai et al., 2009; USAID/Rwanda, 2006).

Applicability of such an approach in other countries

Reliable and accurate routine health data recorded at primary health care facility can provide strategic information for more effective monitoring and evaluation systems in NTD programs in resource-constrained settings (Ledikwe et al., 2014; Mate, Bennett, Mphatswe, Barker, & Rollins, 2009). That is in line with a shared commitment to efforts to combat NTDs initiated by WHO since 2007 (WHO, 2012). Since then, the improvement of data management, data validity and reliability for informed health care planning and decision making through well-structured and operating health management and information systems (HMIS) have been the preoccupation of the health sector and their global partners in all sub-Saharan countries. Nowadays routine health data reporting systems at the district level have been strengthened in most African countries, such as Uganda (Kiberu et al., 2014), Botswana (Ledikwe et al., 2014), Tanzania, Ghana, Mozambique, and Zambia (Mutale et al., 2013). Likewise, an East Africa Policy Forum on health management information systems with support from DFID Health Resource Centre has been initiated ten years ago (DFID/HRC, 2006). Therefore, this study approach using routine health data from high spatial resolution health facilities can only be replicated in other countries if they have a similar density of health facilities.

However, the district level is not small enough for detecting highly focalized schistosomiasis infection hotspots, often with transmission risk areas corresponding to 1km around the houses (Nagi et al., 2014) or 5 km, as maximum distance, around infected wetland sites (Handzel et al., 2003; Hu et al., 2013). The advantage of Rwanda remains on temporal resolution (historical data with high quality and a stable HMIS for more than ten years) and on higher geographical density of health facilities. The administrative setup of the Rwandan health system ensures that HFSA delineation provides an accurate representation of the area from which each health facility will receive its patients. In other countries, the lower geographical density of health facilities implies that HFSAs will extend over larger areas, thereby reducing the sensitivity to detect smaller geographical hotspots of disease transmission. However, that can be fixed if the aspect of the spatial distribution of the population is incorporated in the analysis since most countries have overall lower population density, more cities, and larger rural areas compared to Rwanda. Furthermore, the fraction of infection represented by recorded cases captures the spatial variability. This implies higher incidences in areas (by large service areas or by a large number of served population) with higher prevalence. Then our approach will have added value if used for mapping of disease distribution at a large spatial extent (country scale as a complement to prevalence surveys. It can also help to improve the samples size definition and sampling sites selection for cross-sectional prevalence surveys.

4. Linkages between the spatio-temporal distribution of reclaimed wetlands and schistosomiasis transmission

The research hypothesis was that conversion of wetland for intense agricultural use is associated with increased rates of intestinal helminth infection. Our analysis shows a strong association between intensified agricultural use of wetlands, especially for rice cultivation and schistosomiasis incidence in space and time (throughout Chapter 3 to 6). Wetlands conversion to irrigated areas in Rwanda is rapidly expanding with the aim to secure sufficient agricultural production. About 20,000 hectares of wetland are officially reclaimed and the target is to double it in less than one decade. Irrigation schemes, in turn, are associated with increased transmission of schistosomiasis (Table 16). Utmost wetlands conversion and extending use to rice cropping affects the spatial and temporal distribution pattern of schistosomiasis in Rwanda. This association is also illustrated by the scatter plot of annual rice cropped areas against schistosomiasis cases in Figure 35.

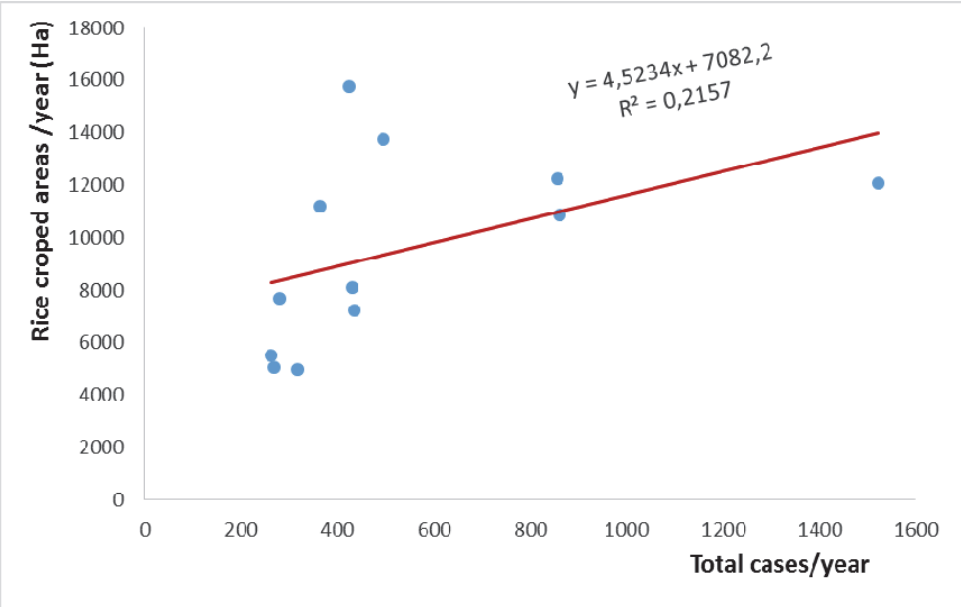

Figure 35: Trend of rice cropped areas (in ha) and schistosomiasis mansoni cases during 2001- 2012

5. Research recommendations

Policy recommendations

Wetland reclamation is seen as an efficient approach of agricultural transformation to ensure food security in Rwanda. This study revealed a location specific sensitivity of wetlands to climate change. Thus, future

planning of wetland development projects should explicitly consider location specific climatic factors as this can help modify wetlands transformation plans for the good.

In concrete terms, our findings confirmed the hypothesis, that wetland conversion into intense agricultural use can both contribute to the well-being of communities, but also increases exposure to helminths resulting in ill health. Therefore, there is a need for an integrated policy of agriculture development and public health. Thus, the challenging trade-off between a higher crop yield (gain money) and disease frequency (costing money) can be attenuated or transformed into synergy by mitigation based on a joint or integrated policy of agriculture development accounting for induced public health problems. A joint action plan combining professionals from the Ministry of Agriculture and Ministry of Natural Resources for an efficient hydrological design for rice cropping (irrigation scheme and dam creation) accompanied by a detailed epidemiological impact assessment and well elaborated mitigation measures in close collaboration with public health professionals from the Ministry of Health and partners from public and private organizations.

Further research areas

Localized understanding and fully quantified relationships between wetland reclamation activities and intestinal helminths are complex. As identified in this research, there is a need for an acute assessment of biophysical, social and economic factors and location specific disease infection counts. Surprisingly, the effect of most available socio-economic factors such as school attendances level, access to improved water for domestic use, proper sanitation was hardly relevant in explaining disease variability at fine scales. Our model outputs did not include socio-economic factors, possibly because the way they are collected shows little variation nation-wide. In Rwanda, the aggregated percentages (administrative district or sector) of school attendance, access to improved water for domestic use, proper sanitation, and wearing shoes are almost homogeneous and reflect the large achievements of the last two decades (MINECOFIN, 2015; United Nations, 2000). Thus, the available data has very limited variation in space for reflecting the detailed conditions created within the areas of wetlands and water development activities. It appears that an extra level of consideration is missing, e.g. the culture of putting off shoes while working in the field, defecation behavior in the bushes around the working place (while far from their home toilet) which may be important for this study is not yet known. Furthermore, it was not possible to explore the monthly or seasonal relationship between disease incidence and risk factors. As illustrated in Figure 36, the annual and seasonal trend of schistosomiasis mansoni exposes

an unstable pattern during last decade in Rwanda. There was some seasonality with peaks which mostly occur in January and July-August.

Location specific model should allow considering the dynamic patterns of wetland extent, rice cropping areas, and schistosomiasis transmission. For instance, a specific model should be developed for small-scale rice cropping in the Central Plateau and another one in the Eastern Province with relatively large scale rice farming into larger wetlands and enormous evapotranspiration from numerous constructed water dams. That prediction should reflect seasonal variation of climatic factors (Muhire & Ahmed, 2016); regular surface runoff and flooding into hierarchical hydrological networks (Munyaneza, Wenninger, & Uhlenbrook, 2012), use of unsafe water (from rivers, dams, and ponds), increasing dry seasons, and induced water born disease transmission (Krauth et al., 2015; Lewin et al., 2007).

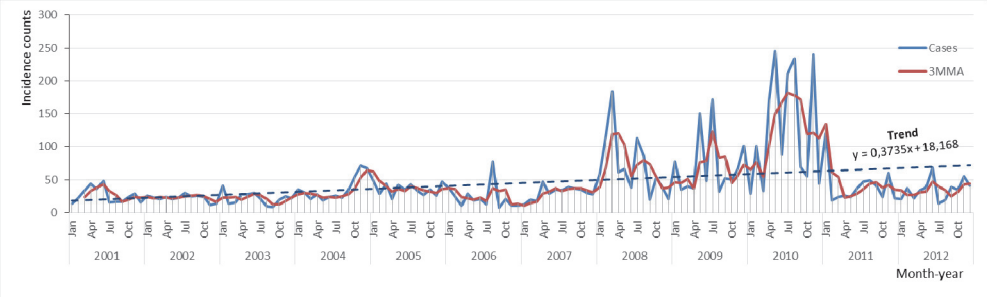

Figure 36: Distribution and seasonal variation of S. mansoni in Rwanda

Therefore, factors that underpin the dynamics of wetland resource use and potential epidemiological impacts should be analyzed using an integrated multi-scale and multi-data approach. Further research may consider using Agent-based modeling (ABM) to capture people behavioral factors, possibly via in-depth interviews, to know what peoples say against what they do.

7.3 A final word

This study generated information that was previously missing, addressing the issue of wetland use and management planning. The detection of marked differences in wetland characteristics at national, agro-ecological zone and watershed level in Rwanda, means that nationwide recommendations should be nuanced. A blanket specific recommendation for wetland use and management plan should be made locally – even at the catchment area level. Routine health data provide valuable information for monitoring the spatial, and spatiotemporal distribution of schistosomiasis. That distribution is very dynamic and routine health data can be efficient to monitor monthly or seasonal variability. The HFSA was proven to be a suitable spatial scale for identification and monitoring of schistosomiasis distribution patterns. The

HFSA is even better for representing the STHs pattern and not the District level, even for prevalence survey.

Climate variation and changes have very significant impact on both the wetlands distribution pattern and the disease distribution pattern. Therefore, detailed climatic data are much more needed than very high spatial resolution DEM for location specific wetlands characterization in Rwanda.

Bibliography

Abera, A., & Nibret, E. (2014). Prevalence of gastrointestinal helminthic infections and associated risk factors among schoolchildren in Tilili town, northwest Ethiopia. *Asian Pacific Journal of Tropical Medicine, 7*(7), 525-530. doi:http://dx.doi.org/10.1016/S1995-7645(14)60088-2

Abera, B., Alem, G., & Melaku, M. (2012). Epidemiology of soil-transmitted helminths, Schistosoma mansoni, and haematocrit values among schoolchildren in Ethiopia. *The Journal of Infection in Developing Countries, 7*(3), 253-260.

Adam, E., Mutanga, O., & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management, 18*(3), 281-296. doi:10.1007/s11273-009-9169-z

Adamas, W. M. (1993). Indigenous use of wetlands and sustainable development in West Africa *The Geographical Journal, Vol 159*(No 2), Pages 209-218.

Albert, D. A., Wilcox, D. A., Ingram, J. W., & Thompson, T. A. (2005). Hydrogeomorphic Classification for Great Lakes Coastal Wetlands. *Journal of Great Lakes Research, 31, Supplement 1*, 129-146. doi:http://dx.doi.org/10.1016/S0380-1330(05)70294-X

Alemayehu, B., & Tomass, Z. (2015). Schistosoma mansoni infection prevalence and associated risk factors among schoolchildren in Demba Girara, Damot Woide District of Wolaita Zone, Southern Ethiopia. *Asian Pacific Journal of Tropical Medicine, 8*(6), 457-463. doi:http://dx.doi.org/10.1016/j.apjtm.2015.05.009

Anselin, L. (1995). The local indicators of spatial association - LISA. *Geographical Analysis, 27*, 93-115. doi:10.1111/j.1538-4632.1995.tb00338.x

Apan, A. A., & Peterson, J. A. (1998). Probing tropical deforestation: The use of GIS and statistical analysis of georeferenced data. *Applied Geography, 18*(2), 137-152. doi:http://dx.doi.org/10.1016/S0143-6228(98)00004-6

Arizono, N., Yamada, M., Tegoshi, T., & Onishi, K. (2012). Molecular Identification of Oesophagostomum and Trichuris Eggs Isolated from Wild Japanese Macaques. *Korean Journal of Parasitology, 50*(3), 253-257. doi:10.3347/kjp.2012.50.3.253

Assare, R. K., Lai, Y. S., Yapi, A., Tian-Bi, Y. N., Ouattara, M., Yao, P. K., . . . N'Goran, E. K. (2015). The spatial distribution of Schistosoma mansoni infection in four regions of western Cote d'Ivoire. *Geospat Health, 10*(1), 345. doi:10.4081/gh.2015.345

Bavia, M. E., Hale, L. F., Malone, J. B., Braud, D. H., & Shane, S. M. (1999). Geographic information systems and the environmental risk of schistosomiasis in Bahia, Brazil. *Am. J. Trop. Med. Hyg., 60*, 566-572.

Beighley, D. H. (2012). Growth and Production of Rice *Soil , Plant Growth and Crop Production, Vol II*.

Bithell, J. F. (2000). A classification of disease mapping methods. *Stat Med, 19*(17-18), 2203-2215.

Boelee, E. (2000). *Irrigation ecology of schistosomiasis : environmental control options in Morocco.* Wageningen University, Wageningen.

Bomblies, A., Duchemin, J.-B., & Eltahir, E. A. B. (2008). Hydrology of malaria: Model development and application to a Sahelian village. *Water Ressources Research, Vol 44*(W12445). doi:10.1029/2008WR006917

Boots, B. N., & Getis, A. (1998). *Point Pattern Analysis*. Newbury Park, CA: Sage Publications.

Brinson, M. M. (1993). *A Hydrogeomorphic Classification for Wetlands* Retrieved from Washington DC 20314- 1000:

Brooker, S., Kabatereine, N. B., Smith, J. L., Mupfasoni, D., Mwanje, M. T., Ndayishimiye, O., . . . Snow, R. W. (2009). An updated atlas of human helminth infections: the example of East Africa. *International Journal of Health Geographics, doi:10.1186/1476-072X-8-42*.

Buis, E., & Veldkamp, A. (2008a). Modelling dynamic water redistribution patterns in arid catchments in the Negev Desert of Israel *Earth Surface Processes and Landforms, 33*, 107-122.

Buis, E., & Veldkamp, A. (2008b). Modelling dynamic water redistribution patterns in arid catchments in the Negev Desert of Israel. *Earth Surface Processes and Landforms, 33*(1), 107-122. doi:10.1002/esp.1531

Burns, R. B., & Burns, R. A. (2009). Logistic Regression *Business Research Methods and Statistics using SPSS* (pp. 565-588): SAGE Publications Ltd.

CDC. (2013). Hookworm Retrieved from http://www.cdc.gov/parasites/hookworm/

CDC. (2014). BioMedware SpaceStat- Empirical Bayesian Smoothing. .

Ceccato, V., & Persson, L. O. (2002). Dynamics of rural areas: an assessment of clusters of employment in Sweden. *Journal of Rural Studies, 18*(1), 49-63. doi:http://dx.doi.org/10.1016/S0743-0167(01)00028-6

Chen, H., Yang, L., Yang, Z. F., & Yu, S. W. (2012). Sustainable reservoir operations to balance upstream human needs and downstream lake ecosystem targets. *Procedia Environmental Sciences, 13*(0), 1444-1457. doi:10.1016/j.proenv.2012.01.136

Chirico, G. B., Western, A. W., Grayson, R. B., & Blöschl, G. (2005). On the definition of the flow width for calculating specific catchment area patterns from gridded elevation data. *Hydrological Processes, 19*(13), 2539-2556. doi:10.1002/hyp.5730

Chitsulo, L., Engels, D., Montresor, A., & Savioli, L. (2000). The global status of schistosomiasis and its control. *Acta Tropica, 77*(1), 41-51. doi:http://dx.doi.org/10.1016/S0001-706X(00)00122-4

Cho, S. Y. (2009). Fifty Years of the Korean Society for Parasitology. *Korean Journal of Parasitology, 47*, S7-S19. doi:10.3347/kjp.2009.47.S.S7

Clements, A., Deville, M.-A., Ndayishimiye, O., Brooker, S., & Fenwick, A. (2010). Spatial co-distribution of neglected tropical diseases in the East African Great Lakes region: revisiting the justification for integrated control. *Tropical Medicine & International Health, 15*(2), 198-207. doi:10.1111/j.1365-3156.2009.02440.x

Clements, A., Firth, S., Dembele, R., Garba, A., Toure, S., Sacko, M., . . . Fenwick, A. (2010). Use of Bayesian geostatistical prediction to estimate local variations in Schistosoma haematobium, infection in Western Africa. *Bull World Health Org, 87*. doi:10.2471/BLT.08.058933

Cliff, A., & Ord, J. (1973). *Spatial Autocorrelation*. London: Pion Limited.

Congdon, P. (2009). A multilevel model for cardiovascular disease prevalence in the US and its application to micro area prevalence estimates. *International Journal of Health Geographics, 8*(1), 1-14. doi:10.1186/1476-072X-8-6

Cowardin, L. M., Carter, V., Golet, F. C., & LaRoe, E. T. (1979). *Classification of Wetlands and Deep Water Habitats of the United States*. Retrieved from Washington, D.C

Cowardin, L. M., Carter, V., Golet, F. C., & LaRose, E. T. (1979). *Classification of Wetlands and Deep Water Habitats of the United States*. Retrieved from Washington, D.C

Creed, I. F., Sanford, S. E., Beall, F. D., Molot, L. A., & Dillon, P. J. (2003). Cryptic wetlands: integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes. *Hydrological Processes, 17*, 3629 - 3648. doi:10.1002/hyp.1357

Curie, F., Gaillard, S., Ducharne, A., & Bendjoudi, H. (2007). Geomorphological methods to characterise wetlands at the scale of the Seine watershed. *Science of The Total Environment, 375*(1–3), 59-68. doi:http://dx.doi.org/10.1016/j.scitotenv.2006.12.013

Currie, B. J. (2001). Environmental Change, Global Warming and Infectious Diseases in Northern Australia. *Environmental Health, Vol 1*(No 4).

Dai, W., & Huang, Y. (2006). Relation of soil organic matter concentration to climate and altitude in zonal soils of China. *CATENA, 65*(1), 87-94. doi:http://dx.doi.org/10.1016/j.catena.2005.10.006

de Koning, G. H. J., Veldkamp, A., & Fresco, L. O. (1998). Land use in Ecuador : a statistical analysis at different aggregation levels. *Agriculture, ecosystems and environment, 70*(2-3), 231-247.

Degallier, N., Favier, C., Menkes, C., Lengaigne, M., Ramalho, W. M., Souza, R., . . . Boulanger, J. P. (2010). Toward an early warning system for

dengue prevention: modeling climate impact on dengue transmission. *Climatic Change, 98*(3-4), 581-592. doi:10.1007/s10584-009-9747-3

Delepierre, G. (1982). Les régions agro-climatiques en relation avec l'intensité de l'erosion du Sol. *Bulletin Agricole du Rwanda, 2*, 87-95.

DFID/HRC. (2006). *East Africa Policy Forum: health management information systems. Plans to improve health information in East Africa*. Retrieved from http://www.eldis.org/go/home&id=23145&type=Document - .WQn2tNykKpr

Dimitriou, E., & Zacharias, I. (2010). Identifying microclimatic, hydrologic and land use impacts on a protected wetland area by using statistical models and GIS techniques. *Mathematical and Computer Modelling, 51*(3–4), 200-205. doi:http://dx.doi.org/10.1016/j.mcm.2009.08.009

Eisenberg, J. N. S., Desai, M. A., Levy, K., Bates, S. J., Liang, S., Naumoff, K., & Scott, J. C. (2007). Environmental Determinants of Infectious Disease: A Framework for Tracking Causal Links and Guiding Public Health Research. *Environ Health Perspect., 115*, 1216 - 1223. doi:10.1289/ehp.9806

Fonseca, F., Freitas, C., Dutra, L., Guimarães, R., & Carvalho, O. (2014). Spatial modeling of the schistosomiasis mansoni in Minas Gerais State, Brazil using spatial regression. *Acta Tropica, 133*(0), 56-63. doi:http://dx.doi.org/10.1016/j.actatropica.2014.01.015

Fontaine, M., Aerts, R., Özkan, K., Mert, A., Gülsoy, S., Süel, H., . . . Muys, B. (2007). Elevation and exposition rather than soil types determine communities and site suitability in Mediterranean mountain forests of southern Anatolia, Turkey. *Forest Ecology and Management, 247*(1–3), 18-25. doi:http://dx.doi.org/10.1016/j.foreco.2007.04.021

Fóti, S., Balogh, J., Nagy, Z., Herbst, M., Pintér, K., Péli, E., . . . Bartha, S. (2014). Soil moisture induced changes on fine-scale spatial pattern of soil respiration in a semi-arid sandy grassland. *Geoderma, 213*(0), 245-254. doi:http://dx.doi.org/10.1016/j.geoderma.2013.08.009

Galletti, C. S., Ridder, E., Falconer, S. E., & Fall, P. L. (2013). Maxent modeling of ancient and modern agricultural terraces in the Troodos foothills, Cyprus. *Applied Geography, 39*(0), 46-56. doi:http://dx.doi.org/10.1016/j.apgeog.2012.11.020

Gatrell, A. C., & Loytonen, M. (1998). *GIS and HEALTH*. Padstow UK: T.J. International Ltd, Padstow.

Gell, P., Mills, K., & Grundell, R. (2013). A legacy of climate and catchment change: the real challenge for wetland management. *Hydrobiologia, 708*(1), 133-144. doi:10.1007/s10750-012-1163-4

Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Anal, 3*, 445 - 450.

Getis, A. (2007). A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis, 40(2008)*, 297-309.

Getis, A., & Ord, J. K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis, 24(3)*, 189-206. doi:10.1111/j.1538-4632.1992.tb00261.x

Goovaerts, P. (2008). Accounting for rate instability and spatial patterns in the boundary analysis of cancer mortality maps. *Environmental and ecological statistics, 15*(4), 421-446. doi:10.1007/s10651-007-0064-6

Grzybowski, K. (1992). The Quaternary formations and the palaeogeomorphology of the Lake Kivu segment of the Central African Rift. *Bulletin - Polish Academy of Sciences: Earth Sciences, 40*(4), 279-300.

Gu, W., & Novak, R. J. (2005). HABITAT-BASED MODELING OF IMPACTS OF MOSQUITO LARVAL INTERVENTIONS ON ENTOMOLOGICAL INOCULATION RATES, INCIDENCE, AND PREVALENCE OF MALARIA. *American Journal of Tropical Medecine and Hygiene, Vol 73*(No 3), pp 546 - 552.

Guimarães, R. J. P. S., Freitas, C. C., Dutra, L. V., Moura, A. C. M., Amaral, R. S., Drummond, S. C., . . . Carvalho, O. S. (2008). Schistosomiasis risk estimation in Minas Gerais State, Brazil, using environmental data and GIS techniques. *Acta Tropica, 108*(2–3), 234-241. doi:http://dx.doi.org/10.1016/j.actatropica.2008.07.001

Halpenny, C. M., Paller, C., Koski, K. G., Valdés, V. E., & Scott, M. E. (2013). Regional, Household and Individual Factors that Influence Soil Transmitted Helminth Reinfection Dynamics in Preschool Children from Rural Indigenous Panamá. *PLoS Neglected Tropical Diseases, 7*(2), e2070. doi:10.1371/journal.pntd.0002070

Handzel, T., Karanja, D. M. S., Addiss, D. G., Hightower, A. W., Rosen, D. H., Colley, D. G., . . . Secor, W. E. (2003). Geographic distribution of schistosomiasis and soil-transmitted helminths in Western Kenya: Implications for anthelminthic mass treatment. *American Journal of Tropical Medicine and Hygiene, 69*(3), 318-323.

Hanotier, J., & Gigase, P. L. (1981). Note on a new focus of schistosomiasis (S. mansoni) in Rwanda. *Ann Soc Belg Med Trop, 61*(1), 93-98.

Hansen, M. K., Brown, D. J., Dennison, P. E., Graves, S. A., & Bricklemyer, R. S. (2009). Inductively mapping expert-derived soil-landscape units within dambo wetland catenae using multispectral and topographic data. *Geoderma, 150*(1–2), 72-84. doi:http://dx.doi.org/10.1016/j.geoderma.2009.01.013

Heady, C. (1992). Values for the environment: A guide to economic appraisal. J. T. Winpenny HMSO, London, 1991, 277 pp. *Public Administration and Development, 12*(5), 512-512. doi:10.1002/pad.4230120509

Hengl, T., Gruber, S., & Shrestha, D. P. (2003). Digital Terrain Analysis in Ilwis

Hessami, M., Gachon, P., Ouarda, T. B. M. J., & St-Hilaire, A. (2008). Automated regression-based statistical downscaling tool. *Environmental Modelling & Software, 23*(6), 813-834. doi:http://dx.doi.org/10.1016/j.envsoft.2007.10.004

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very Hig Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology, 25*, 1965-1978. doi:10.1002/joc.1276

Hogg, A. R., & Todd, K. W. (2007). Automated discrimination of upland and wetland using terrain derivatives. *Can. J. Remote Sensing, 33*(1), S68 - S83.

Hollis, G. E. A., W.M.; Aminu-Kano. (1993). The Hadejia-Nguru Wetlands: Environment, economy and sustainable development of a Sahelian floodplain wetland *CABI Book* (pp. 244 pages).

Hotez, P. J. (2014). Ten Global "Hotspots" for the Neglected Tropical Diseases. *PLoS Neglected Tropical Diseases, 8*(5), e2496. doi:10.1371/journal.pntd.0002496

Hu, Y., Ward, M. P., Xia, C., Li, R., Sun, L., Lynn, H., . . . Jiang, Q. (2016). Monitoring schistosomiasis risk in East China over space and time using a Bayesian hierarchical modeling approach. *Scientific Reports, 6*, 24173. doi:10.1038/srep24173

http://www.nature.com/articles/srep24173 - supplementary-information

Hu, Y., Zhang, Z., Chen, Y., Wang, Z., Gao, J., Tao, B., . . . Jiang, Q. (2013). Spatial pattern of schistosomiasis in Xingzi, Jiangxi Province, China: the effects of environmental factors. *Parasites & Vectors, 6*(1), 214. doi:10.1186/1756-3305-6-214

Hurlimann, E., Schur, N., Boutsika, K., Stensgaard, A., Laserna de Himpsl, M., Ziegelbauer, K., . . . Vounatsou, P. (2011). Toward an open-access global database for mapping, control, and surveillance of neglected tropical diseases. *PLoS Negl Trop Dis, 5*, e1404.

Isabwe, A., Ruberanziza, E., Mupfasoni, D., Ruxin, J., Clerinx, J., & White, P. (2012). Potential for Transmission of Schistosomiasis in Kayonza District *Rwanda Medical Journal, 69*(2), 14-19.

Jalloh, A., Roy-Macauley, H., & Sereme, P. (2012). Major agro-ecosystems of West and Central Africa: Brief description, species richness, management, environmental limitations and concerns. *Agriculture, Ecosystems & Environment, 157*(0), 5-16. doi:10.1016/j.agee.2011.11.019

Kabatereine, N. B., Brooker, S., Tukahebwa, E. M., Kazibwe, F., & Onapa, A. W. (2004). Epidemiology and geography of Schistosoma mansoni in Uganda: implications for planning control. *Tropical Medicine & International Health, 9*(3), 372-380. doi:10.1046/j.1365-3156.2003.01176.x

Kabo, G. J., Paulechka, Y. U., Zaitsau, D. H., & Firaha, A. S. (2015). Prediction of the enthalpies of vaporization for room-temperature ionic liquids: Correlations and a substitution-based additive scheme. *Thermochimica Acta, 609*, 7-19. doi:http://dx.doi.org/10.1016/j.tca.2015.04.013

Kaliappan, S. P., George, S., Francis, M. R., Kattula, D., Sarkar, R., Minz, S., . . . Kang, G. (2013). Prevalence and clustering of soil-transmitted helminth infections in a tribal area in southern India. *Tropical Medicine & International Health, 18*(12), 1452-1462. doi:10.1111/tmi.12205

Karim, F., Kinsey-Henderson, A., Wallace, J., Arthington, A. H., & Pearson, R. G. (2012). Modelling wetland connectivity during overbank flooding in a tropical floodplain in north Queensland, Australia. *Hydrological Processes, 26*(18), 2710-2723. doi:10.1002/hyp.8364

Karuranga, E. (2011). INTESTINAL SCHISTOSOMIASIS AND STH STOOL SCREENING EXERCISE IN NKOMBO ISLAND, RUSIZI DISTRICT. *RBC/TRAC Plus EID Unit, Unipublished Report*.

Keddy, P. (2010). *Wetland Ecology: Principles and Conservation* Cambridge,UK.

Kharkar, M., & Bowalekar, S. (2014). Extent of Under Reporting of Adverse Drug Reactions (ADRs) in India: Evaluation using Logistic Regression Analysis (LRA) Model. *Clinical Trials, 4*(1). doi:http://dx.doi.org/10.4172/2167-0870.1000155

Kiberu, V. M., Matovu, J. K., Makumbi, F., Kyozira, C., Mukooyo, E., & Wanyenze, R. K. (2014). Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. *BMC Medical Informatics and Decision Making, 14*(1), 40. doi:10.1186/1472-6947-14-40

Knight, M., Elhelu, O., Smith, M., Haugen, B., Miller, A., Raghavan, N., . . . Brindley, P. J. (2015). Susceptibility of Snails to Infection with Schistosomes is influenced by Temperature and Expression of Heat Shock Proteins. *Epidemiology (Sunnyvale), 5*(2). doi:10.4172/2161-1165.1000189

Kok, K., & Veldkamp, A. (2001). Evaluating impact of spatial scales on land use pattern analysis in Central America *Agriculture, ecosystems and environment, 85*(1-3), 205-221.

Kok, K., & Veldkamp, T. (2011). Scale and Governance: Conceptual Considerations and Practical Implications. *Ecology and Society, 16*(2).

Krauth, S. J., Musard, C., Traore, S. I., Zinsstag, J., Achi, L. Y., N'Goran, E. K., & Utzinger, J. (2015). Access to, and use of, water by populations living in a schistosomiasis and fascioliasis co-endemic area of northern Cote d'Ivoire. *Acta Tropica, 149*, 179-185. doi:10.1016/j.actatropica.2015.05.019

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and Methods, 26*(6), 1481-1496. doi:10.1080/03610929708831995

Kulldorff, M. (2015). Software for the spatial and space-time scan statistics. (Version SaTScan (TM) V9.4.2): http://www.satscan.org/.

Lai, Y. S., Zhou, X. N., Utzinger, J., & Vounatsou, P. (2013). Bayesian geostatistical modelling of soil-transmitted helminth survey data in the People's Republic of China. *Parasites & Vectors, 6*, 16. doi:10.1186/1756-3305-6-359

Large, A. R. G., & Petts, G. E. (1996). Historical channel-floodplain dynamics along the River Trent: Implications for river rehabilitation. *Applied Geography, 16*(3), 191-209. doi:http://dx.doi.org/10.1016/0143-6228(96)00004-5

Larsen, M. N., & Roepstorff, A. (1999). Seasonal variation in development and survival of Ascaris suum and Trichuris suis eggs on pastures. *Parasitology, 119*, 209-220. doi:10.1017/s0031182099004503

Law, J., & Quick, M. (2013). Exploring links between juvenile offenders and social disorganization at a large map scale: a Bayesian spatial modeling approach. *Journal of Geographical Systems, 15*(1), 89-113. doi:10.1007/s10109-012-0164-1

Ledikwe, J. H., Grignon, J., Lebelonyane, R., Ludick, S., Matshediso, E., Sento, B. W., . . . Semo, B.-w. (2014). Improving the quality of health information: a qualitative assessment of data management and reporting systems in Botswana. *Health Research Policy and Systems, 12*(1), 7. doi:10.1186/1478-4505-12-7

Lehmann, E. L. (1998). *Nonparametrics : statistical methods based on ranks* (Revised first edition ed.). Upper Saddle River: Prentice Hall.

Lemeshow, S., & Moeschberger, M. L. (2005). Review of Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models by Vittinghoff, Glidden, Shiboski, and McCulloch. *The Strata, 5*(2), 274-278.

Leuchowius, K. (2014). *Report on the Health Care Sector and Business Opportunities in Rwanda*. Retrieved from http://www.swecare.se/Portals/swecare/Documents/Report-on-the-Health-Care-Sector-and-Business-Opportunities-in-Rwanda-Sep2014-vers2.pdf

Lewin, S., Norman, R., Narman, N., Thomas, E., Bradshaw, D., & South African Comparative Risk, A. (2007). Estimating the burden of disease attributable to unsafe water and lack of sanitation and hygiene in South Africa in 2000. *Samj South African Medical Journal, 97*(8), 755-762.

Lichtenberg, E. (2002). Chapter 23 Agriculture and the environment. In L. G. Bruce & C. R. Gordon (Eds.), *Handbook of Agricultural Economics* (Vol. Volume 2, Part A, pp. 1249-1313): Elsevier.

Limpert, E., stahel, W. A., & Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues. *BioScience, 51*(5).

Long, J. A., Nelson, T. A., & Wulder, M. A. (2010). Characterizing forest fragmentation: Distinguishing change in composition from configuration. *Applied Geography, 30*(3), 426-435. doi:http://dx.doi.org/10.1016/j.apgeog.2009.12.002

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Murray, C. J. L. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*(9859), 2095-2128. doi:http://dx.doi.org/10.1016/S0140-6736(12)61728-0

Macchioni, F., Segundo, H., Gabrielli, S., Totino, V., Gonzales, P. R., Salazar, E., . . . Cancrini, G. (2015). Dramatic Decrease in Prevalence of Soil-Transmitted Helminths and New Insights Into Intestinal Protozoa in Children Living in the Chaco Region, Bolivia. *American Journal of Tropical Medicine and Hygiene, 92*(4), 794-796. doi:10.4269/ajtmh.14-0039

Malesu, M. M., Oduor, A. R., Chrogony, K., Nyolei, D., Gachene, C. K. K., Biamah, E. K., . . . Mogoi, J. (2010). Rwanda Irrigation Master Plan. The Government of Rwanda, Ministry of Agriculture and Animal Resources, Ebony Company Limited and World Agroforestry Centre (ICRAF). Nairobi, Kenya. Retrieved from http://www.worldagroforestry.org/downloads/Publications/PDFS/B16738.pdf

Mate, K. S., Bennett, B., Mphatswe, W., Barker, P., & Rollins, N. (2009). Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *Plos One, 4*. doi:10.1371/journal.pone.0005483

McNab, W. H. (1989). Terrain Shape Index: Quantifying Effect of Minor Landforms on Tree Height. *Forest Science, 35*(1), 91-104.

McSweeney, R. (2011). *Rwanda's Climate: Observations and Projections*. Retrieved from United Kingdom:

Melloh, R. A., Racine, C. H., Sprecher, S. W., Greeley, N. H., & Weyrick, P. B. (1999). *Comparisons of Digital Terrain Data for Wetland Inventory on Two Alaskan Army Bases*. Retrieved from Hanover, N.H. Special Report 99-15. 21 pp.:

Mendoza-Sanchez, I., Phanikumar, M. S., Niu, J., Masoner, J. R., Cozzarelli, I. M., & McGuire, J. T. (2013). Quantifying wetland–aquifer interactions in a humid subtropical climate region: An integrated approach. *Journal of Hydrology, 498*(0), 237-253. doi:http://dx.doi.org/10.1016/j.jhydrol.2013.06.022

Meurs, L., Mbow, M., Boon, N., van den Broeck, F., Vereecken, K., Dièye, T. N., . . . Polman, K. (2013). Micro-Geographical Heterogeneity in Schistosoma mansoni and S. haematobium Infection and Morbidity in a

Co-Endemic Community in Northern Senegal. *PLoS Negl Trop Dis, 7*(12), e2608. doi:10.1371/journal.pntd.0002608

MINAGRI. (2013). *Rwanda's Rice Commodity Chain - Facing Globalisation* Retrieved from Kigali:

MINAGRI, & ICRAF. (2010). *Rwanda Irrigation Master Plan.* : World Agroforestry Centre (ICRAF), World Agroforestry Centre (ICRAF) Water Management Unit, Ebony Enterprises Ltd (Israel), Ministry of Agriculture and Animal Resources (MINAGRI).

MINALOC. (2005). *Organic Law No 29/2005 of 31/12/2005 Determining the Administrative Entities of the Republic of Rwanda*. Kigali: Official Gazette of the Republic Rwanda Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/5067/4160.pdf?sequence=1.

MINALOC. (2008). *Rwanda Decentralization Implementation Program; Ministry of Local Governance, Kigali.* .

MINECOFIN. (2015). *Demographic and Health Survey [DHS] 2014/2015. Key findings*. Kigali.

Ministry of Agriculture and Animal Resources. (2004). Strategic Plan for Agricultural Transformation in Rwanda; Ministry of Agriculture and Animal Resources of Rwanda

Mitsch, W. J., & Gosselink, J. G. (2000). THE VALUES OF WETLANDS: LANDSCAPES AND INSTITUTIONAL PERSPECTIVES. The value of wetlands: importance of scale and landscape setting. *Ecological Economics, 35(200)*(Special Issue), 25-33.

Montresor, A., Crompton, D. W. T., Hall, A., Bundy, D. A. P., & Savioli, L. (1998). *Guidelines for Evaluation of Soil-Transmitted Helminthiasis and Schistosomiasis at Community Level*. Retrieved from Geneva: http://whqlibdoc.who.int/hq/1998/who_ctd_sip_98.1.pdf

Moore, D. A., & Carpenter, T. E. (1999). Spatial Analytical Methods and Geographic Information Systems: Use in Health Research and Epidemiology. *Epidemiologic Reviews, 21*(2). doi:10.1093/oxfordjournals.epirev.a017993

Mote, K. E., Makanga, B., & Kisakye, J. J. M. (2005). Prevalence of intestinal parasites among school children in Moyo district, Uganda. *Health Policy and Development*(volume 3 number 2), 184-186.

Muhire, I., & Ahmed, F. (2015). Spatio-temporal trend analysis of precipitation data over Rwanda. *South African Geographical Journal, 97*(1), 50-68. doi:10.1080/03736245.2014.924869

Muhire, I., & Ahmed, F. (2016). Spatiotemporal trends in mean temperatures and aridity index over Rwanda. *Theoretical and Applied Climatology, 123*(1-2), 399-414. doi:10.1007/s00704-014-1353-2

Mukashema, A., Veldkamp, A., & Vrieling, A. (2014). Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network. *International Journal of Applied Earth Observation and*

Geoinformation : JAG, 33, 331-340. doi:http://dx.doi.org/10.1016/j.jag.2014.05.005

Munyaneza, O., Wenninger, J., & Uhlenbrook, S. (2012). Identification of runoff generation processes using hydrometric and tracer methods in a meso-scale catchment in Rwanda. *Hydrol. Earth Syst. Sci., 16*(7), 1991-2004. doi:10.5194/hess-16-1991-2012

Mupfasoni, D., Karibushi, B., Koukounari, A., Ruberanziza, E., Kaberuka, T., Kramer, M. H., . . . Fenwick, A. (2009). Polyparasite Helminth Infections and Their Association to Anaemia and Undernutrition in Northern Rwanda.

Mutale, W., Chintu, N., Amoroso, C., Awoonor-Williams, K., Phillips, J., Baynes, C., . . . Sherr, K. (2013). Improving health information systems for decision making across five sub-Saharan African countries: Implementation strategies from the African Health Initiative. *BMC Health Services Research, 13*(Suppl 2), S9-S9. doi:10.1186/1472-6963-13-S2-S9

Mutuku, F. M., Bayoh, M. N., Hightower, A. W., Vulule, J., Gimnig, J. E., Mueke, J. M., . . . Walker, E. D. (2009). A supervised land cover classification of a western Kenya lowland endemic for human malaria: associations of land cover with larval Anopheles habitats. *International Journal of Health Geographics, Vol 8*(19). doi:10.1186/1476-072X-8-19

Nabahungu, N. L. (2012). *Problems and opportunities of wetland management in Rwanda.* (PhD), Wageningen University Wageningen.

Nabahungu, N. L., & Visser, S. M. (2011). Farmer's Knowledge and Perception of Agricultural Wetland Management in Rwanda. *Land Degradation & Development*, doi: 10.1002/ldr.1133. doi:10.1002/ldr.1133

Nagi, S., Chadeka, E. A., Sunahara, T., Mutungi, F., Justin, Y. K., Kaneko, S., . . . Hamano, S. (2014). Risk factors and spatial distribution of Schistosoma mansoni infection among primary school children in Mbita District, Western Kenya. *PLoS Negl Trop Dis, 8*(7), e2991. doi:10.1371/journal.pntd.0002991

Nagumo, N., Sugai, T., & Kubo, S. (2013). Late Quaternary floodplain development along the Stung Sen River in the Lower Mekong Basin, Cambodia. *Geomorphology, 198*(0), 84-95. doi:http://dx.doi.org/10.1016/j.geomorph.2013.05.016

National Institute of Statistics of Rwanda. (2012). *Rwanda Demographic and Health Survey 2010.* Retrieved from Calverton, Maryland, USA:

Ndayambaje, J., Heijman, W. J. M., & Mohren, G. M. J. (2012). Household Determinants of Tree Planting on Farms in Rural Rwanda. *Small-scale Forestry*.

NISR. (2010a). *Rwanda Demographic and Health Survey 2010. Final Report; Kigali.* .

NISR. (2010b). *Rwanda statistical yearbook 2009*. Retrieved from

NISR. (2014). *Fourth Rwanda Population and Housing Census. Thematic Report: Population Projections*. Kigali.

Nkurikiye, J. B. (2016). Rice Cooperatives in Rwanda: A Perspective on Resource Use. . *Imperial Journal of Interdisciplinary Research, 2*(2).

Nuvolone, D., della Maggiore, R., Maio, S., Fresco, R., Baldacci, S., Carrozzi, L., . . . Viegi, G. (2011). Geographical information system and environmental epidemiology: a cross-sectional spatial analysis of the effects of traffic-related air pollution on population respiratory health. *Environmental Health, 10*. doi:10.1186/1476-069x-10-12

Nyandwi, E., Veldkamp, A., Osei, F. B., & Amer, S. (2017). Spatio-temporal dynamics of schistosomiasis in Rwanda between 2001 and 2012 : impact of the national Neglected Tropical Disease control programme. *Geospatial health : health applications in geospatial science : open access, 12*(1), 11 p.

Nyandwi, E., Veldkamp, T., & Amer, S. (2016). Regional climate sensitivity of wetland environments in Rwanda: the need for a location-specific approach. *Regional Environmental Change, 16*(6), 1635-1647. doi:10.1007/s10113-015-0905-z

Nyandwi, E., Veldkamp, T. A., & Amer, S. (2015). Regional climate sensitivity of wetland environments in Rwanda: the need for a location-specific approach. *Regional Environmental Change, 15*(8). doi: 10.1007/s10113-015-0905-z

Nzigidahera, B. (2007). *Study of the Flora in the Cohoha sub-basin; Wetland and Biodiversity Component of Nile Transboundary Environmental Action Project/UNOPS*. Retrieved from

Odoi, A., Martin, S. W., Michel, P., Holt, J., Middleton, D., & Wilson, J. (2003). Geographical and temporal distribution of human giardiasis in Ontario, Canada. *Int J Health Geogr, 2*(1), 5. doi:10.1186/1476-072x-2-5

Ord, J. K., & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis, 27(4)*, 286-306. doi:10.1111/j.1538-4632.1995.tb00912.x

Ortega Hinojosa, A. M., Davies, M. M., Jarjour, S., Burnett, R. T., Mann, J. K., Hughes, E., . . . Jerrett, M. (2014). Developing small-area predictions for smoking and obesity prevalence in the United States for use in Environmental Public Health Tracking. *Environmental Research, 134*(0), 435-452. doi:http://dx.doi.org/10.1016/j.envres.2014.07.029

Osei, F. (2014). Current Statistical Methods for Spatial Epidemiology: A Review. *Austin Biometrics and Biostatistics, 1*(2).

Osei, F., & Duker, A. (2008). Spatial and demographic patterns of Cholera in Ashanti region - Ghana. *International Journal of Health Geographics, 7*(1), 44.

Paez, A., & Scott, D. (2004). Spatial statistics for urban analysis: a review of techniques with examples. *GeoJournal, 61*. doi:10.1007/s10708-005-0877-5

Phillips, D. R. (1990). *Health and health care in the third world*. Harlow: Longman Scientific & Technical.

Prioul, C., & Sirven, P. (Cartographer). (1981). Atlas du Rwanda; Kigali-Paris-Nantes

Pullan, R. L., Kabatereine, N. B., Quinnell, R. J., & Brooker, S. (2010). Spatial and Genetic Epidemiology of Hookworm in a Rural Community in Uganda. *PLoS Negl Trop Dis, 4*(6), e713. doi:10.1371/journal.pntd.0000713

Ralph, T. J., & Hesse, P. P. (2010). Downstream hydrogeomorphic changes along the Macquarie River, southeastern Australia, leading to channel breakdown and floodplain wetlands. *Geomorphology, 118*(1–2), 48-64. doi:http://dx.doi.org/10.1016/j.geomorph.2009.12.007

Ramsar. (1971). *Wetland Disease Manual: Guidelines for Assessment, Monitoring and Management of Animal Disease in Wetlands.* Retrieved from http://www.ramsar.org/pdf/lib/rtr7-disease.pdf

Ramsar. (2007). Wetlands International. Ramsar Sites Database (accessed 15.08.2012).

Ravi, R. P., & Kulasekaran, R. A. (2013). Trends in Reproductive Tract Infections and Barriers to Seeking Treatment among Young Women: A Community Based Cross Sectional Study in South India. *American Journal of Epidemiology and Infectious Disease, 1*(4), 53-58.

REMA. (2008). *Etablissement d'un inventaire national rapide des marais et élaboration de 5 avant-projets d'arrêtés ministériels relatifs aux marais* Retrieved from Kigali http://www.rema.gov.rw/rema_doc/publications/Rapport de l'Inventaire des Marais du Rwanda Module 3.pdf

REMA. (2009). *Rwanda State of Environment and Outlook. Our Environment for Economic Development* Retrieved from Kigali: http://www.unep.org/pdf/rwanda_outlook.pdf

Rochlin, I., Turbow, D., Gomez, F., Ninivaggi, D. V., & Campbell, S. R. (2011). Predictive Mapping of Human Risk for West Nile Virus (WNV) Based on Environmental and Socioeconomic Factors. *Plos One, 6*(8). doi:10.1371/journal.pone.0023280

Rollemberg, C. V., Silva, M. M., Rollemberg, K. C., Amorim, F. R., Lessa, N. M., Santos, M. D., . . . Jesus, A. R. (2015). Predicting frequency distribution and influence of sociodemographic and behavioral risk factors of Schistosoma mansoni infection and analysis of co-infection with intestinal parasites. *Geospat Health, 10*(1), 303. doi:10.4081/gh.2015.303

Rothman, K. J. (2012). *Epidemiology: An Introduction.* (2nd Edition ed.). Oxford: Oxford University Press, Inc.

Ruberanziza, E., Kabera, M., Ortu, G., Kanobana, K., Mupfasoni, D., Ruxin, J., . . . Katja, P. (2015). Nkombo Island: The most important Schistosomiasis mansoni focus in Rwanda *American Journal of Life Sciences, 3*(1), 27-31. doi:10.11648/j.ajls.20150301.16

Ruberanziza, E., Mupfasoni, D., Karibushi, B., Kabera, M., Karema, C., Nyatanyi, T., . . . Ruxin, J. (2010). A Recent Update of Schistosomiasis Mansoni Endemicity Around Lake Rweru. *Rwanda Medical Journal/Revue Médicale Rwandaise, 68*(4).

Ruxin, J., & Negin, J. (2012). Removing the neglect from neglected tropical diseases: The Rwandan experience 2008–2010. *Global Public Health: An International Journal for Research, Policy and Practice., 7*(8), 812-822. doi:10.1080/17441692.2012.699535

Rwanda, R. o. (2010). *Enabling Self Sufficiency and Competitiveness of Rwanda Rice. Issues and Policy Options*. Kigali Retrieved from http://www.minagri.gov.rw/fileadmin/user_upload/documents/agridocs/Rwa_RicePolicyReport.pdf.

Sabin Vaccine Institute. (2011). *Working toward sustainable, country-led programs the Burundi and Rwanda. NTD control programs year three annual report May 2009 - May 2010*. Schistosomiasis Control Initiative, CBM, The Earth Institute, Geneva Global

Sacko, M., Magnussen, P., Keita, A. D., Traoré, M. S., Landouré, A., Doucouré, A., . . . Vennervald, B. J. (2011). Impact of Schistosoma haematobium infection on urinary tract pathology, nutritional status and anaemia in school-aged children in two different endemic areas of the Niger River Basin, Mali. *Acta Tropica, 120, Supplement 1*(0), S142-S150. doi:10.1016/j.actatropica.2010.12.009

Safari, B. (2012). Trend Analysis of the Mean Annual Temperature in Rwanda during the Last Fifty Two Years. *Journal of Environmental Protection, 3*(6), 538-551. doi: 10.4236/jep.2012.36065

Schmidlin, T., Hürlimann, E., Silué, K. D., Yapi, R. B., Houngbedji, C., Kouadio, B. A., . . . Raso, G. (2013). Effects of Hygiene and Defecation Behavior on Helminths and Intestinal Protozoa Infections in Taabo, Côte d'Ivoire. *Plos One, 8*(6), e65722. doi:10.1371/journal.pone.0065722

Schulte, L., Veit, H., Burjachs, F., & Julià, R. (2009). Lütschine fan delta response to climate variability and land use in the Bernese Alps during the last 2400 years. *Geomorphology, 108*(1–2), 107-121. doi:http://dx.doi.org/10.1016/j.geomorph.2007.11.014

Schur, N., Hurlimann, E., Stensgaard, A.-S., Chimfwembe, K., Mushinge, G., Simoonga, C., . . . Vounatsou, P. (2013). Spatially explicit Schistosoma infection risk in eastern Africa using Bayesian geostatistical modelling. *Acta Tropica, 128*(2013), 365-377.

Simoonga, C., Kazembe, L. N., Kristensen, T. K., Olsen, A., Appleton, C. C., Mubita, P., & Mubila, L. (2008). The epidemiology and small-scale

spatial heterogeneity of urinary schistosomiasis in Lusaka province, Zambia. *Geospat Health, 3*(1), 57-67. doi:10.4081/gh.2008.232

Standley, C. J., Adriko, M., Alinaitwe, M., Kazibwe, F., Kabatereine, N. B., & Stothard, J. R. (2009). Intestinal schistosomiasis and soil-transmitted helminthiasis in Ugandan schoolchildren: a rapid mapping assessment. *Geospatial Health, 4*(1), 39-53.

Steinmann, P., Keiser, J., Bos, R., Tanner, M., & Utzinger, J. (2006). Schistosomiasis and Water Resources Development: Systematic Review, Meta-Analysis, and Estimates of People at Risk. *The Lancet Infectious Diseases, 6*. doi:10.1016/S1473-3099(06)70521-7

Strunz, E. C., Addiss, D. G., Stocks, M. E., Ogden, S., Utzinger, J., & Freeman, M. C. (2014). Water, Sanitation, Hygiene, and Soil-Transmitted Helminth Infection: A Systematic Review and Meta-Analysis. *PLoS Medicine, 11*(3), e1001620. doi:10.1371/journal.pmed.1001620

Swedesurvey. (2010). *Project of Rwanda National Land Use and Development Master Plan, Report for Production of ortho photo in Rwanda.* Rwanda Natural Resources Authority- Land and Mapping Department. Kigali

Swetnam, R. D., & Reyers, B. (2011). Meeting the challenge of conserving Africa's biodiversity: The role of GIS, now and in the future. *Landscape and Urban Planning, 100*(4), 411-414. doi:10.1016/j.landurbplan.2011.02.002

Tango, T. (2010). *Statistical Methods for Disease Clustering. Springer*.

Temme, A. J. A. M., Schoorl, J. M., & Veldkamp, A. (2006). Algorithm for dealing with depressions in dynamic landscape evolution models *Computers & Geosciences, 32*, 452-461.

Thaxton, M. (2009). *Integrating Population, Health, and Environment in Rwanda*. Retrieved from

Thomas, S. M., Fischer, D., Fleischmann, S., Bittner, T., & Beierkuhnlein, C. (2011). RISK ASSESSMENT OF DENGUE VIRUS AMPLIFICATION IN EUROPE BASED ON SPATIO-TEMPORAL HIGH RESOLUTION CLIMATE CHANGE PROJECTIONS. *Erdkunde, 65*(2), 137-150.

TRAC+. (2008). *National Prevalence Survey on Soil-Transmitted Helminths and Schistosomiasis in School -Aged Children* TRAC PLUS - Centre for Infectious Disease Control /Access Project

Tsai, P.-J., Lin, M.-L., Chu, C.-M., & Perng, C.-H. (2009). Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006. *BMC Public Health, 9*(1), 1-13. doi:10.1186/1471-2458-9-464

United Nations. (2000). *Millenium Declaration* Retrieved from http://www.un.org/en/development/devagenda/millennium.shtml

USAID/Rwanda. (2006). *Rwanda HMIS Assessment Report*. Retrieved from Kigali http://pdf.usaid.gov/pdf_docs/PNADG504.pdf

Utzinger, J., Raso, G., Brooker, S., De Savigny, D., Tanner, M., ØRnbjerg, N., . . . N'Goran, E. K. (2009). Schistosomiasis and neglected tropical

diseases: towards integrated and sustainable control and a word of caution. *Parasitology, 136*(13), 1859-1874. doi:10.1017/S0031182009991600

van der Werf, M. J. (2003). *Schistosomiasis morbidity and management of cases in Africa.* (Ph. D), Erasmus MC, University Medical Center, Rotterdam.

Van Oost, K., Govers, G., Cerdan, O., Thauré, D., Van Rompaey, A., Steegen, A., . . . Poesen, J. (2005). Spatially distributed data for erosion model calibration and validation: The Ganspoel and Kinderveld datasets. *CATENA, 61*(2–3), 105-121. doi:http://dx.doi.org/10.1016/j.catena.2005.03.001

Vanek, M. J., Shoo, B., Mtasiwa, D., Kiama, M., Lindsay, S. W., Fillinger, U., . . . Killeen, G. F. (2006). Community-based surveillance of malaria vector larval habitats: a baseline study in urban Dar es Salaam, Tanzania. *BMC Public Health, Vol 6*(154). doi:10.1186/1471-2458-6-154

Vega-Corredor, M. C., & Opadeyi, J. (2014). Hydrology and public health: linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies. *Earth Perspectives, 1*(1), 1-14. doi:10.1186/2194-6434-1-3

Veldkamp, A., & Fresco, L. O. (1997). Reconstructing land use drivers and their spatial scale dependence for Costa Rica (1973 and 1984). *Agricultural Systems, 55*(1), 19-43. doi:http://dx.doi.org/10.1016/S0308-521X(95)00080-O

Verdoodt, A., & van Ranst, E. (2006). Environmental assessment tools for multi-scale land resources information systems: A case study of Rwanda. *Agriculture, Ecosystems & Environment, 114*(2–4), 170-184. doi:http://dx.doi.org/10.1016/j.agee.2005.10.006

Vergne, T., Paul, M. C., Chaengprachak, W., Durand, B., Gilbert, M., Dufour, B., . . . Grosbois, V. (2014). Zero-inflated models for identifying disease risk factors when case detection is imperfect: Application to highly pathogenic avian influenza H5N1 in Thailand. *Preventive Veterinary Medicine, 114*(1), 28-36. doi:http://dx.doi.org/10.1016/j.prevetmed.2014.01.011

Walz, Y., Wegmann, M., Dech, S., Raso, G., & Utzinger, J. (2015). Risk profiling of schistosomiasis using remote sensing: approaches, challenges and outlook. *Parasit Vectors, 8*, 163. doi:10.1186/s13071-015-0732-6

Wang, Z. J., Chapiro, J., Schernthaner, R., Duran, R., Chen, R. X., Geschwind, J. F., & Lin, M. D. (2015). Multimodality 3D Tumor Segmentation in HCC Patients Treated with TACE. *Academic Radiology, 22*(7), 840-845. doi:10.1016/j.acra.2015.03.001

WHO. (2003). *Controlling disease due to helminth infections*. Geneva.

WHO. (2006). *Preventive chemotherapy in human helminthiasis: coordinated use of anthelminthic drugs in control interventions : a manual for health professionals and program managers*. Retrieved from Geneva:

WHO. (2010a). *Rwanda - Country Profile. Preventive Chemotherapy and Transmission Control.* Retrieved from Department of Control of Neglected Tropical Diseases, Geneva/ Switzerland: http://www.who.int/neglected_diseases/preventive_chemotherapy/databank/CP_Rwanda.pdf

WHO. (2010b). *Schistosomiasis, Fact Sheet No 115; February 2010. World Health Organization. Available at* http://www.who.int/mediacentre/factsheets/fs115/en/*. Accessed: Oct 31, 2016.* Retrieved from

WHO. (2012). *Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases. A Roadmap for Implementation.* : Geneva: World Health Organization.

Wu, J., Liu, M., & Davis, J. (2005). Land consolidation and productivity in Chinese household crop production. *China Economic Review, 16*(1), 28-49. doi:http://dx.doi.org/10.1016/j.chieco.2004.06.010

Wu, J., Wang, J., Meng, B., Chen, G., Pang, L., Song, X., . . . Zheng, X. (2004). Exploratory spatial data analysis for the identification of risk factors to birth defects. *BMC Public Health, 4*, 23-23. doi:10.1186/1471-2458-4-23

Xie, Z., Liu, Z., Jones, J. W., Higer, A. L., & Telis, P. A. (2011). Landscape unit based digital elevation model development for the freshwater wetlands within the Arthur C. Marshall Loxahatchee National Wildlife Refuge, Southeastern Florida. *Applied Geography, 31*(2), 401-412. doi:10.1016/j.apgeog.2010.10.003

Yamazaki, D., Baugh, C. A., Bates, P. D., Kanae, S., Alsdorf, D. E., & Oki, T. (2012). Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling *Journal of Hydrology, 436-437*, 81-91.

Zhao, X., Stein, A., & Chen, X.-L. (2011). Monitoring the dynamics of wetland inundation by random sets on multi-temporal images. *Remote Sensing of Environment, 115*(9), 2390-2401. doi:http://dx.doi.org/10.1016/j.rse.2011.05.002

Zhou, D., Gong, H., & Liu, Z. (2008). Integrated ecological assessment of biophysical wetland habitat in water catchments: Linking hydro-ecological modelling with geo-information techniques. *Ecological Modelling, 214*(2–4), 411-420. doi:http://dx.doi.org/10.1016/j.ecolmodel.2008.03.014

Summary

Wetlands comprise the transitional zone between land and water with land covered by shallow water or with a water table at or near the surface, and an ecosystem associated with long-term soil inundation. Worldwide, wetlands sustain large communities of people, who depend on their natural resources that maintain them. Also in Rwanda, wetlands are ecosystems of critical importance to almost 12 million people, generally engaged in self-sufficient smallholder agriculture. Since the 1970's agricultural development programs have been implemented in inland valley swamps, larger inland deltas and lacustrine wetlands in Western and Eastern Africa. In Rwanda, the conversion of wetlands for intensified agricultural production is a more recent phenomenon. In the framework of Vision 2020, the long-term strategy for national development, and the Economic Development and Poverty Reduction Strategy (EDPRS), rural development and agricultural transformation are spearheads for rapid and sustainable development. Wetland conversion is one of the major mechanisms to increase agricultural production, ensure food security for the growing population, and achieve MDG1 (eradicating extreme poverty and hunger). Apart from benefits, the "created irrigation scheme landscape" can also have negative impacts. One of these is the increased risk of transmission of schistosomiasis and soil-transmitted helminths, which can negatively affect the health status of communities that depend on their produce. This dissertation contributes to improved understanding of the linkages between wetlands and wetland use on the one hand, and transmission risk of schistosomiasis and Soil Transmitted Helminths (STH) on the other.

The *first* stage of the study modelled and mapped wetlands and their characteristics both at a national and regional scale. Findings show that the interactions between geomorphologic setting (elevation, slope), hydrological conditions (contributing area) and climatic factors (temperature, rainfall) are fundamental for characterizing Rwandan wetlands. Logistic regression analysis was conducted to generate wetland probability maps with very high accuracy both at national and subnational level. Further analysis was then done to estimate in how far Rwandan wetlands are sensitive to climate change. Results indicate that, countrywide, a 1% increase in temperature would cause a net probable wetland area decline of 12.4%. A 1% increase in rainfall would cause a net probable wetland increase of 1.6%. The analysis also showed that wetlands in the central part of Rwanda are the most sensitive to climate change. Geographic variability of climate change effects implies that future wetlands use planning should be location-specific.

The *second* stage of the research analysed the linkages between environmental (including wetlands and wetland use) and socio-economic risk

factors on the one hand, and transmission risk of schistosomiasis and STHs on the other. The novelty of our analysis is that it is based upon routinely collected data of confirmed cases recorded at primary health facility level. Confirmed case data were aggregated to health facility service areas (HFSAs) with a high spatial resolution, and merged with population data to generate incidence rates. Compared to traditional District-level prevalence surveys, our fine-grained spatial resolution enables a geographically much more accurate detection of transmission hotspots and associated risk factors. Outcomes clearly show that the spatial pattern of schistosomiasis incidence has a localised distribution and is therefore better monitored at HFSA then at the much larger District level. Regarding risk factors, a significant positive relation was identified between schistosomiasis incidence and presence of and proximity to rice cropped areas, rainfall and specific soil properties. For STHs incidence, population density and the proportion of rural residents emerged as significantly associated risk factors.

The *third* research stage consisted of a spatiotemporal cluster analysis using SaTScan statistics. Findings demonstrate that areas of high schistosomiasis transmission vary considerably across space and over time. The results show a clear impact of the national intervention program, but not leading to full eradication. The analysis revealed existence of three types of clusters: areas with persistently high rates of transmission (close to lakes), areas where transmission declines over time (impact of NTD program), and areas of high transmission that emerge over time. Newly emerging clusters are associated with increased agricultural use of wetlands, irrigated rice cultivation in particular. Our findings illustrate that the spatial and temporal distribution of schistosomiasis incidence is very dynamic which makes elimination difficult to achieve. Identified spatio-temporal dynamics should, therefore, be considered in future disease control activities.

The *fourth* stage of this study focused on forecasting the future spatial pattern of schistosomiasis transmission risk using a Bayesian model that accounts for false zero cases. The zero-inflated Poisson model incorporated anticipated climate–induced changes in rainfall and temperature as well as the planned expansion of rice cultivation in wetlands. Other factors were assumed to be constant. Results show that, changes between relative risk of 2009 and forecasted risk in 2050 identify persisting and emerging areas with high relative risk of schistosomiasis transmission. Model outcomes indicate that the risk of schistosomiasis transmission can be expected to be 69% higher in areas with rice cultivation, 29% higher in areas in close proximity to rice farming, and 50% higher in areas close to water bodies. The prediction and forecasting maps provide a valuable tool for monitoring schistosomiasis risk in Rwanda and planning future disease control initiatives.

Overall, this research has confirmed that the anticipated impacts of climate change on wetlands will not be uniform across Rwanda but will vary from place to place. Wetlands in the central part of Rwanda are the most sensitive to climate change. In addition, the research confirmed that the anticipated future spatiotemporal expansion of rice cropping areas is positively correlated with increased risk of transmission of schistosomiasis. Finally, routinely collected health data are proven to be a valuable information source for monitoring the spatial, and spatiotemporal distribution of schistosomiasis. That distribution is very dynamic and routine health data can be efficient to monitor monthly or seasonal variability. Conducting the analysis at the level of the health facility service area, as the basic spatial unit of analysis, proved to be a suitable spatial scale for modelling, monitoring and mapping of schistosomiasis transmission across space and over time.

Samenvatting

Wetlands vormen de overgangszone tussen land en water met land bedekt door ondiep water, een grondwaterspiegel aan of nabij het oppervlak, en een ecosysteem dat samenhangt met langdurige bodeminundatie. Wereldwijd zijn grote delen van de bevolking afhankelijk van de natuurlijke hulpbronnen die wetlands bieden. Ook in Rwanda zijn wetlands van cruciaal belang voor de bijna 12 miljoen mensen die voor het merendeel werkzaam zijn in de zelfvoorzienende kleinschalige landbouw. Sinds de jaren zeventig van de vorige eeuw zijn ontwikkelingsprogramma's in West- en Oost-Afrika uitgevoerd om moerassen, binnenlandse delta's en laaggelegen gebieden langs rivieren te benutten voor landbouwdoeleinden. In Rwanda is het gebruik van wetlands voor intensievere vormen van landbouw echter een recenter verschijnsel.

In het kader van Vision 2020, het lange termijnplan voor de nationale ontwikkeling van Rwanda, en de strategie voor economische ontwikkeling en armoedebestrijding (EDPRS), zijn plattelandsontwikkeling en transformatie van de agrarische sector beleidsspeerpunten om te komen tot snelle en duurzame economische ontwikkeling. Het gebruik van wetlands voor landbouwdoeleinden wordt daarbij gezien als een van de belangrijkste manieren om de landbouwproductie te verhogen, de voedselzekerheid voor de groeiende bevolking te waarborgen en SDG2 (honger uitbannen, zorgen voor voedselzekerheid en duurzame landbouw) te verwezenlijken. Behalve voordelen heeft het gebruik van wetlands voor geïntensiveerde landbouw ook potentieel negatieve gevolgen. Eén daarvan is een verhoogd risico op besmetting door schistosomiasis en geo-helminthen, met name voor bevolkingsgroepen werkzaam in voor landbouw gebruikte wetlands. Doel van dit proefschrift is het verkrijgen van een beter inzicht in de relaties tussen het gebruik van wetlands voor landbouwdoeleinden enerzijds en risico's van verspreiding en transmissie van schistosomiasis en geo-helminthen anderzijds.

De eerste fase van het onderzoek bestond uit het modelleren en in kaart brengen van wetlands en hun kenmerken op zowel nationaal als regionaal schaalniveau. Onderzoeksresultaten wijzen uit dat de interacties tussen geomorfologie (elevatie, helling), hydrologische omstandigheden (stroomgebied) en klimatologische factoren (temperatuur, regenval) van fundamenteel belang zijn voor het karakteriseren van Rwandese wetlands. Met behulp van logistische regressieanalyses op zowel nationaal als sub-nationaal niveau zijn kaarten gegenereerd die met hoge nauwkeurigheid wetlands afbakenen. Vervolgens is nader geanalyseerd in hoeverre Rwandese wetlands gevoelig zijn voor klimaatverandering. Resultaten geven aan dat een stijging van de temperatuur met 1% in Rwanda als geheel zal

leiden tot een daling van 12.4% van het nationale wetland-areaal. Een toename van de neerslag met 1% zal leiden tot een uitbreiding van het wetland-areaal met 1.6%. Uit de analyses werd eveneens duidelijk dat wetlands in het centrale deel van Rwanda het meest gevoelig zijn voor klimaatverandering. Het feit dat de verwachte effecten van klimaatverandering in Rwanda van gebied tot gebied zullen variëren impliceert dat de planning van het toekomstige gebruik van wetlands voor landbouwdoeleinden locatie-specifiek moet zijn.

De tweede fase van het onderzoek richtte zich op het analyseren van de relaties tussen het gebruik van wetlands, een selectie van milieu- en sociaaleconomische factoren enerzijds, en het besmettingsrisico van schistosomiasis en geo-helminthen anderzijds. Innovatief aan de ontwikkelde methodiek is dat deze is gebaseerd op routinematig verzamelde gegevens van geïnfecteerde patiënten, voor elke individuele gezondheidskliniek afzonderlijk. Voor elk van de ongeveer 375 klinieken is vervolgens het verzorgingsgebied in kaart gebracht. Incidentie van schistosomiasis en geo-helminthen per verzorgingsgebied werd daarna berekend door bevolkingsgegevens en het aantal geïnfecteerde patiënten te combineren. Een belangrijk voordeel van deze methodiek is dat verzorgingsgebieden ruimtelijk fijnmazig zijn. In de standaard prevalentieonderzoeken in Rwanda, daarentegen worden uitkomsten geaggregeerd naar het districtsniveau (n= 30) welke geografisch veel grotere gebieden beslaan. De hoge ruimtelijke resolutie van onze methodiek maakt het mogelijk om hotspots van schistosomiasis transmissie en daaraan gerelateerde risicofactoren geografisch veel nauwkeuriger te detecteren. Uit het onderzoek blijkt dat schistosomiasis incidentie ruimtelijk zeer gedifferentieerd is en daarom beter geanalyseerd en gemonitord kan worden op het niveau van verzorgingsgebied dan op het veel grotere districtsniveau. Met betrekking tot risicofactoren werd een significant positief verband vastgesteld tussen de incidentie van schistosomiasis en (nabijheid van) rijstteeltgebieden, (hogere) neerslag en specifieke bodemeigenschappen. Incidentie van geo-helminthen bleek positief geassocieerd met bevolkingsdichtheid en het aandeel rurale bevolking in een verzorgingsgebied maar vertoonde geen uitgesproken ruimtelijke variatie.

De derde onderzoeksfase bestond uit een ruimtelijk-temporele clusteranalyse op basis van SaTScan-statistieken. Uitkomsten tonen aan dat gebieden met verhoogde schistosomiasistransmissie een aanzienlijke geografische en temporele variatie vertonen. Het nationale interventie programma, gestart in 2007, heeft weliswaar geresulteerd in een duidelijke reductie van schistosomiasis incidentie in bepaalde gebieden maar het heeft zeker niet tot volledige eliminatie geleid. De analyses tonen het bestaan aan van drie types geografische clusters: gebieden met een aanhoudend hoge transmissie (bij

meren), gebieden waar de transmissie in de loop der tijd afneemt (effect van het interventieprogramma), en gebieden waar in de loop der tijd een verhoogde transmissie optreedt. Deze opkomende clusters zijn aantoonbaar geassocieerd met het toegenomen gebruik van wetlands voor landbouwdoeleinden, met name voor geïrrigeerde rijstteelt. Toekomstige zorginterventies zouden hier bij voorkeur rekening mee moeten houden.

De vierde fase van deze studie richtte zich op het voorspellen van het toekomstige ruimtelijke patroon van schistosomiasistransmissie op basis van Bayesiaanse Poisson regressies met een zero-inflated component. Verwachte klimaatveranderingen (hogere neerslag en temperatuur) en de geplande uitbreiding van rijstcultivatie in wetlands werden in het model opgenomen, de overige factoren werden als constant beschouwd. Gemodelleerde resultaten voor 2050 identificeren niet alleen gebieden met een aanhoudend hoge transmissie maar ook nieuwe gebieden waar in de loop der tijd een verhoogd risico op transmissie kan worden verwacht. De resultaten van de modellen geven aan dat het toekomstig risico op schistosomiasistransmissie naar verwachting 69% hoger zal zijn in gebieden met rijstteelt, 29% hoger in gebieden in de nabijheid van de rijstteelt en 50% hoger in gebieden in de buurt van waterlichamen. De gegenereerde, ruimtelijk-specifieke, prognoses vormen een waardevol instrument voor het monitoren van schistosomiasis in Rwanda en kunnen worden gebruikt ter ondersteuning van toekomstige zorginterventies.

In het algemeen heeft dit onderzoek bevestigd dat de verwachte effecten van klimaatverandering op wetlands in Rwanda niet uniform zullen zijn, maar van plaats tot plaats zullen verschillen. Wetlands in het centrale deel van Rwanda zijn het meest gevoelig voor klimaatverandering. Ook bevestigde het onderzoek dat (de verwachte) uitbreiding van rijstcultivatie in wetlands positief gecorreleerd is met een verhoogd risico op schistosomiasis.

Tenslotte heeft het onderzoek aangetoond dat het gebruik van routinematig verzamelde zorggegevens een waardevolle aanvullende bron van informatie zijn voor identificatie, analyse en monitoren van de ruimtelijke en ruimtelijk-temporele verspreiding van schistosomiasis in Rwanda. Het gebruik van het verzorgingsgebied als ruimtelijke eenheid voor de analyses bleek eveneens zeer geschikt. De hoge ruimtelijke en temporele resolutie van dit type zorggegevens op verzorgingsgebied niveau maakt het tevens mogelijk om maandelijkse of seizoensgebonden variaties te monitoren.

Biography

Elias Nyandwi was born on 05th July 1971 in Ruhango, Rwanda. In 2003, he graduated from the National University of Rwanda where he obtained a bachelor degree in Human and Physical Geography. He pursued his post-graduate studies at the International Institute of Geo-Information Science and Earth Observation (ITC), The Netherlands and graduated in March 2008 with a degree of Master of Science in Geo-Information Science and Earth Observation for Environmental System Analysis and Management. Since July 2012 he was awarded the NUFFIC Scholarship, under the NICHE/RWA/071 project, to pursue his doctoral research at the Faculty of Geo-Information Science and Earth Observation of the University of Twente (ITC-UT), the Netherlands. His research outputs were presented and published in high profile (regional and international) conferences and journals and resulted in this thesis.

Elias was respectively employed at National Service of Census (SNR) until March 2005 and at the Centre for Geographic Information System and Remote Sensing of the University of Rwanda (CGIS – UR) as an Assistant Researcher, since June 2005 up to now. He was involved in several research projects, teaching and tailor-made training, consultancy and community service related to GIS and RS and applications. He also worked as Head of Environmental and Natural Resources Management research unit and National Coordinator of three years (2009-2012) Pan-African Research Project on Participatory GIS for Forest Resources Management funded by International Development Research Centre (IDRC).

Scientific publications

Nyandwi, E., Veldkamp, A., Frank Osei, F., and Amer, S (2017). Spatiotemporal dynamics of Schistosomiasis mansoni incidences rates in Rwanda between 2001 and 2012: impact of the national NTD control program. Geospatial health, 12(1), 48-58.

Nyandwi, E., Veldkamp, A., and Amer, S (2016). Regional climate sensitivity of wetland environments in Rwanda: the need for a location - specific approach. Regional environmental change, 16(6), 1635-1647.

Nduwamungu J., **Nyandwi, E.,** Mazimpaka J.D, Mugiraneza T., Mukashema A., Uwayezu E., Rwanyiziri G. and Nzabanita V (2013). Rwanda Forest Cover Mapping Using High-Resolution Aerial Photographs. In: Proceedings of Global

Geospatial Data Infrastructure GSDI 14, 4-8 November 2013, Addis Abeba, Ethiopia.

Nyandwi, E. and MUKASHEMA, A. (2011). Excessive Deforestation of Gishwati Mountainous Forest and Biodiversity changes, e-bulletin of Enda Lead Africa ict4d Research Program in March 2011.

Nyandwi, E., (2010). GIS and RS for Modeling Landform Controls on the Distribution of Forest Canopy Vascular Epiphyte. International Journal of Ecology & Development, Volume 16(10), 2010.

International Conferences

Nyandwi, E., Veldkamp, A., Amer, S., Karema, C. and Umulisa, I. Spatial patterns of soil-transmitted helminth types and associated environmental risk factors in Rwanda. 12th European Multicolloqium of Parasitology, 20–24 July 2016, Turku, Finland.

Nyandwi, E., Veldkamp, A., Amer, S., Karema, C. and Umulisa, I. Fine Resolution Mapping of *Schistosomiasis mansoni* incidence and risk factors. 12th European Multi-colloquim of Parasitology, 20–24 July 2016, Turku, Finland.

Nyandwi, E., Veldkamp, A. and Amer, S. Delineating Rwandan Wetland using Topographic and Climatic parameters. International Conference on Geospatial Technologies for Sustainable Urban and Rural Development (GeoTech Rwanda), 18-20 November 2015, Kigali/Rwanda.

Nyandwi, E., Veldkamp, A., and Amer, S. Wetland conversion and Exposure to schistosomiasis and intestinal helminths. 4th Annual East African Health and Scientific Conference; International Health Exhibition and Trade Fair. 27-29 March 2013, Kigali, Rwanda.

Nyandwi, E. Landform Controls on the Distribution of Forest Canopy Vascular Epiphyte in a Tropical Mountainous Rainforest using GIS and RS: Case Study of Nyungwe National Park, South-Western Rwanda. 28th Annual Esri User Conference, 4-8 August 2008, San Diego, California.

ITC Dissertation List

https://www.itc.nl/Pub/research_programme/Research-review-and-output/PhD-Graduates