Increasing capacity in disease vector modelling to improve malaria and arbovirus mitigation strategies

By: Rachelle Desrochers

Report Type: Final Technical Progress Report

Date: April 30, 2015

Full Name of Research Institution: HealthBridge and Kilimanjaro Christian Medical University College (KCMUC)

Address of Research Institution:
       HealthBridge: 1 Nicholas Street, Suite 1004, Ottawa, Ontario K1N 7B7
       KCMUC: P.O. Box 2240, Moshi, Tanzania

Name(s) of Researchers/Members of Research Team and Contact Infomation:
       Dr. Rachelle Desrochers, GeoHealth Specialist, HealthBridge – rdesrochers@healthbridge.ca
       Professor Franklin Mosha, Professor and Director of Research and Consultancies, KCMUC – fwmosha@gmail.com
       Dr. Manisha Kulkarni, Assistant Professor, Department of Epidemiology and Community Medicine University of Ottawa – manisha.kulkarni@uottawa.ca
       Dr. Jeremy Kerr, Full Professor, Department of Biology, University of Ottawa – jkerr@uottawa.ca
       Dr. Michael Mahande, Post-doctoral researcher, KCMUC – jmmahande@gmail.com
       Filemon Tenu, PhD Candidate, KCMUC – filemonitenu@hotmail.com
       Debora Charles Kajeguka, PhD Candidate, KCMUC – dkajeguka@gmail.com
       Robert Kaaya, Research Assistant, KCMUC – robertkaaya@yahoo.com

Abstract:

Malaria is the leading cause of child mortality in sub-Saharan Africa (World Malaria Report 2012). Malaria case management is difficult, particularly in areas where there are other illnesses such as arbovirus infections (such as chikungunya and dengue fever) also cause fevers and contribute to malaria misdiagnosis. Monitoring disease vector (mosquito) habitat is key in distinguishing between occurrences of arboviral and malarial infections and therefore an important obstacle to the improved management of both diseases. Species distribution modelling, using MAXENT, offers a promising avenue for monitoring disease vector habitat. The long term goal of this project was to improve the capacity of researchers at KCMUC and NIMR in Tanzania in species distribution modelling that can improve vector control strategies. Through workshops and remote collaboration, we worked together to develop maps of where suitable habitat is found for the mosquito vectors of malaria and the vectors of arboviruses, using this as a teaching case to transfer skills in species distribution modelling and spatial statistics. Our research partners now have access to ArcGIS and have gained useful skills in GIS and species distribution modelling. Moreover, a productive partnership has developed between KCMUC and NIMR with the University of Ottawa and HealthBridge.

Keywords:

Synthesis: *The report should begin with a half-page to one-page synthesis of the project which situates the work in the period with respect to the project as a whole. The abstract of the project written when the grant was initially approved by IDRC and the objectives listed in the grant agreement should be useful inputs when preparing this part of the report.*

Malaria is the leading cause of child mortality in sub-Saharan Africa (World Malaria Report 2012). Malaria case management is difficult, particularly in areas where there are other illnesses such as arboviruses that contribute greatly to malaria misdiagnosis and overtreatment. Monitoring disease vector (mosquito) habitat is key in distinguishing between occurrences of arboviral and malarial infections and therefore an important obstacle to the improved management of both diseases. Species distribution modelling, using Maximum Entropy modelling, offers a promising avenue for monitoring disease vector habitat. The long term goal of this project is to improve the capacity of researchers at Kilimanjaro Christian Medical University College (KCMUC) in Tanzania in species distribution modelling that can improve vector control strategies.

The long term goal of this project is to increase the in-house capacity of KCMUC and the National Institute for Medical Research (NIMR) for research in vector dynamics and vector control strategies and, ultimately, reduce the prevalence of malaria and arboviruses. We are using a project that our Tanzanian partners have identified as being of interest (mapping suitable habitat for mosquitoes that transmit malaria and those that transmit arboviruses) as a teaching case to transfer skills in species distribution modelling and spatial statistics. Together, we are developing maps of suitable vector habitat (*Anopheles sp.* and *Aedes sp.*) that will be validated against historic data of vector occurrence and disease prevalence to confirm the key environmental factors that govern the distribution of suitable habitat and thereby disease risk. The resulting maps of suitable habitat for the mosquitoes that transmit malaria relative to those that transmit arbovirus infections as well as improved understanding of the relationships between vector habitat and disease risk will be invaluable to vector control strategies and management of these diseases.

The secondary long-term goal of this project is to strengthen the budding research partnership between KCMUC (including NIMR) with the University of Ottawa and HealthBridge that began in 2009 through the development of the University of Ottawa's field course on the Ecology of East African Ecosystems by Dr. Manisha Kulkarni, which included collaboration with malaria researchers at KCMUC.

The intermediate and immediate objectives are:
1. To identify environmental factors that govern the spatial prevalence of malaria relative to arboviral infections.
a. To generate a map of suitable habitat for *Anopheles* mosquitoes (malaria vectors)
b. To generate a map of suitable habitat for *Aedes* mosquitoes (arbovirus vectors)
2. To identify probable hotspots for malaria and arbovirus infections
3. To increase capacity of researchers to carry out high resolution vector modelling and conduct spatial statistical analyses.

The specific goals of the project were:

G1. A first workshop to provide advanced training in GIS and Maximum Entropy (maxent) modelling (a form of species distribution modelling to generate suitable habitat maps) to our research collaborators (Filemon Tenu – PhD candidate; Debora Kajeguka– PhD candidate; Michael Mahande – Post-doctoral fellow) and colleagues at KCMUC.

G2. To generate preliminary habitat maps of the mosquito vectors, led by Mr. Tenu, Ms. Kajeguka and Dr. Mahande (intermediate and immediate objective 1a and 1b above).

G3. To conduct field work to validate the preliminary habitat maps, led by Mr.Tenu, Ms. Kajeguka and Dr. Mahande (intermediate and immediate objective 1a and 1b above).

G4. A second workshop to provide training in backcasting to validate the species distribution models as well as spatial statistical methods to relate clinical incidences of malaria and arbovirus infections with the spatial distributions of suitable habitat for the disease specific vector.

G5. A final workshop to discuss the results and plan the preparation of manuscripts presenting the results.

The research problem: *The reader should be reminded of the basic rationale of the project and the research problem or problems being addressed. Often, the researchers' understanding of the problems will have evolved since the project was approved. The report should describe this evolution and the reasons behind it. Restate the objectives where necessary. Major changes must be acknowledged and agreed to by IDRC.*

The key research steps in this project are the development of maps of suitable vector habitat (*Anopheles sp.* and *Aedes sp*.) that will be validated against historic data of vector occurrence and disease prevalence to confirm the key environmental factors that govern the distribution of suitable habitat and thereby disease risk and hotspots. We are focusing on three specific areas of Tanzania: Muleba district in the northwest of Tanzania near Rwanda and Burundi; Moshi and Hai districts in northern Tanzania at the base on Mount Kilimanjaro; and Muheza district in northeast of Tanzania along the coast of the Indian Ocean.

The initial aim of the research component of the project was to take advantage of existing records (recent and historic) of mosquito occurrences and malaria prevalence, using the recent mosquito occurrence records to generate suitable habitat maps. We planned to backcast the habitat models (project them back in time 10 years) and test them against the historic mosquito records. Habitat models that accurately capture the important drivers of species distributions should accurately predict historic occurrences when projected back in time.

We have modified our strategy for developing and validating the vector habitat maps for three reasons: A) challenges in assembling the existing mosquito occurrence records (both recent and historic), B) an unseasonably dry warm season and C) the preference expressed by a reviewer of the original grant application for field validation of the habitat maps.

 A)   Challenges in assembling the existing mosquito occurrence records:
The species distribution models generate maps of habitat suitability by relating species occurrence points to maps of environmental variables. Species occurrence points are records of which species (in this case, of mosquito) was captured (or observed) along with the geographic coordinates of the location where it was captured and, when possible, the date of capture. Occurrence points of *Anopheles sp*. and *Aedes sp*. mosquitoes have been collected for over a decade as part of numerous studies conducted by KCMUC, NIMR and other Tanzanian organizations with whom they collaborate. Unfortunately, these records often occur only in paper form and/or the geographic coordinate system in which the coordinates were measured is not documented (geographic coordinates are generally taken in latitude and longitude but can also be taken in a variety of other coordinate systems such as UTM). In such cases, inclusion of these data in our study involves not only requesting permission to use the data but also that someone be available to enter the data electronically and provide information on the coordinate system used to take the geographic coordinates. Consequently, we were unable to assemble previously collected occurrence points prior to the first workshop and have shifted our focus to capitalize on the mosquito occurrence data currently being collected by Mr Tenu for his PhD thesis in the district of Muheza and the mosquito occurrence data currently being collected by Ms Kajeguka for her PhD thesis in the districts of Moshi and Hai.

We developed four sets of habitat maps based on the following data sources. 1) Tanzania-wide maps using data from an on-going study on the development of insecticide resistance in *Anopheles arabiensis* and *An. gambiae ss* mosquitoes (those that transmit malaria). *Aedes sp.* mosquitoes were not collected as part of this study therefore only maps of the *Anopheles sp.* could be generated. These data are collected annually, during the wet season, from sites throughout the country. We have used the data from 2011 to model suitable

anopheline habitat. They were used serve as a comparison for the more local maps being developed for Muleba, Muheza and Moshi/Hai districts.

Data were available from a 2 year study (2011-2012) longitudinal study in Muleba, where mosquitoes were captured monthly from households in 40 villages throughout the district in 2011(Protopopoff *et al*, Malaria Journal, 2013). These data were used to generate maps of suitable habitat for *An. arabiensis*, *An. gambiae ss* and *Aedes sp*.

Two of our research partners, Filemon Tenu and Debora Kajeguka are undertaking doctoral research through KCMUC and/or NIMR. For their theses, Mr Tenu is conducting extensive mosquito sampling throughout Muheza district and Ms Kajeguka is sampling mosquitoes in Moshi and Hai districts. Both of these studies are producing larger and more complete datasets of mosquito occurrences in the regions than have been previously available. Previous studies in Muheza included only a few sites focusing on resistance studies and focused on *Anopheles sp.* (no previous collection of *Aedes sp.*) while the previous study in Moshi/Hai districts sampled *Aedes sp.* as well as *Anopheles sp.* but only at a handful of sites. Some data collection took place in Muheza in February 2014 and maps of suitable habitat for *Aedes sp*. and *Anopheles sp*. were produced. We hoped to produce individual maps for *Anopheles gambiae ss* and *Anopheles arabiensis* but the mosquito samples have not yet been tested to identify the individual species. Once the laboratory differentiation of species has taken place, Mr Tenu will generate habitat suitability maps for the individual species. The maps for Muheza will be validated through data collection for a vector control intervention taking place throughout the district. The project experienced numerous delays and, consequently, it was not possible to conduct independent field validation of these maps in the time frame of the current grant.

Data collection was delayed in Moshi and Hai districts due to unseasonably dry weather resulting in too few mosquitoes to be sampled and low incidences of malaria and other fevers in local health clinics. Ms Kajeguka, with the assistance of Robert Kaaya – a research assistant at KCMUC who participated in the workshops, collected mosquitoes from sites in the catchment area of the Boma Ng'ombe health facility in Hai and data on patients with malaria, chikungunya and dengue fever from the health facility itself in the summer of 2014. Additional mosquito collections took place in the areas surrounding KCMUC in October 2014 and throughout rural and urban Moshi in January and February 2015. Laboratory analyses for species identification and testing mosquitoes for infection with malaria or an arbovirus is on-going. Ms Kajeguka will use the GIS and species modelling skills acquired through the workshops and our continuing collaboration to create maps of dry season habitat for *Aedes sp.* and *Anopheles sp* for her doctoral thesis*.* She will use these maps to assess dry season risk of malaria and arboviruses. Her thesis also includes data collection in the current wet season. Our collaboration allowed her not only to incorporate vector habitat modelling to her thesis but also to extend its scope to include a comparison of dry season and wet season malaria and arbovirus risk factors.

Given the preference expressed by a reviewer of the initial grant proposal for field validation of the habitat maps rather than backcasting, we have chosen, where possible, to focus grant funds on field work for validating the preliminary habitat maps. Muleba district, in the northwest of Tanzania, is very isolated from KCMUC (in Moshi, Tanzania). The logistical cost of conducting sampling to field validate the habitat maps would be prohibitive therefore we made the decision to focus our resources on validating the maps for Muheza district and Moshi/Hai districts. Researchers at KCMUC are currently conducting a new study in Muleba, where mosquitoes will be collected throughout the previous study area and extending beyond it. We have obtained permission to use these data to validate the habitat suitability maps for Muleba once they are available. Some mosquito occurrence data from 2012 are available for Moshi district from a previous Master's student, Lucille Lyaruu. Ms Kajeguka will use these to test the accuracy of the Moshi portion of the maps as part of her doctoral thesis. The Canadian graduate student at the University of Ottawa, Ms Emily Acheson, was unable to use backcasting to validate Tanzania-wide habitat models for *An. arabiensis* and *An. gambiae s.s.* as part of her MSc thesis work due to insufficient country-wide mosquito records. This has highlighted an important gap in knowledge for malaria prevention in Tanzania and is informing our future research directions.

Research findings: *The main research results to date should be described and interpreted by highlighting the contribution to knowledge that this project represents from a scientific and policy perspective.*

The research outputs can be grouped into three categories: mosquito occurrence records assembled, environmental data assembled and processed, and maps generated.

Table 1: Number and source of occurrence records assembled by species and study area.

Study area	Species available	Data sources	Notes
Tanzania	*Anopheles arabiensis* and *Anopheles gambiae ss*	10 records for *Anopheles arabiensis* and 3 records for *Anopheles gambiae s.s.* for 2010-2011 from the on-going resistance study conducted by NIMR.	
		1 record selected at random for each of *Anopheles arabiensis* and *Anopheles gambiae ss* for Muleba district in 2011 from the longitudinal study described in Protopopoff *et al*., Malaria Journal, 2013.	A random was assigned to the records for 2011. The occurrence point with the smallest random number for each of *Anopheles arabiensis* and *Anopheles gambiae s.s.* was selected.
		1 record selected at random for each of *Anopheles arabiensis* and *Anopheles gambiae ss* from the data collected by Mr Tenu.	A random was assigned to the records. The occurrence point with the smallest random number for each of *Anopheles arabiensis* and *Anopheles gambiae s.s.* was selected.
Muleba district	*Anopheles arabiensis, Anopheles gambiae ss* and *Aedes sp.*	From the longitudinal study described in Protopopoff *et al*., Malaria Journal, 2013: 138 records for *Anopheles arabiensis* for 2011; 428 records for *Anopheles gambiae s.s.*; 19 records for *Aedes aegypti* for 2011 and 2012 (there were only 5 records for 2011).	
Muheza district	*Anopheles arabiensis, Anopheles gambiae ss* and *Aedes sp.*	Data collected within the context of Mr Tenu's PhD thesis: 96 occurrence records for *Anopheles sp*; 22 occurrence records for *Aedes sp*.	Laboratory analyses are currently being conducted to differentiate between *Anopheles arabiensis* and *Anopheles gambiae ss.*
Moshi/Hai districts	*Anopheles arabiensis, Anopheles gambiae ss* and *Aedes sp*	Data have been collected from sites throughout the districts of Hai and Moshi. Laboratory analyses, funded through a Grand Challenges Canada grant to HealthBridge, can be conducted to determine the species.	

Table 2: Environmental and demographic data used for the species distribution modelling with data description and source.

Data set	Data description	Source
Bioclim data	The Bioclim data layers from Worldclim are taken from monthly temperature and rainfall data to create variables that have more biological meaning and are often useful for species distribution modelling. The variables represent annual trends (such as mean annual temperature), seasonality (such as annual range in temperature), and extreme or limiting environmental factors (such as temperature of the coldest month).	We used the ESRI grids for 30 arc-seconds (~1 km) data for the *Bioclim* layers http://www.worldclim.org/current
Elevation	Elevation data was taken from the Shuttle Radar Topography Mission (SRTM) data. The value of each pixel is elevation in meters.	The data are available for download at the following website: http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1#download.
Land cover for 2011	Land cover data derived from the MODIS Terra sensor for 2011. The land cover is classified using the IGBP global vegetation classification scheme.	More information on these data can be found at the following website: https://lpdaac.usgs.gov/products/modis_products_table/mcd12q1
Human population density for 2011	Human population density data were obtained from the Oak Ridges National Laboratory through the Geographic, Statistical and Government Information Centre of the University of Ottawa library.	More information on these data and sources are available at the following website: http://web.ornl.gov/sci/landscan/index.shtml

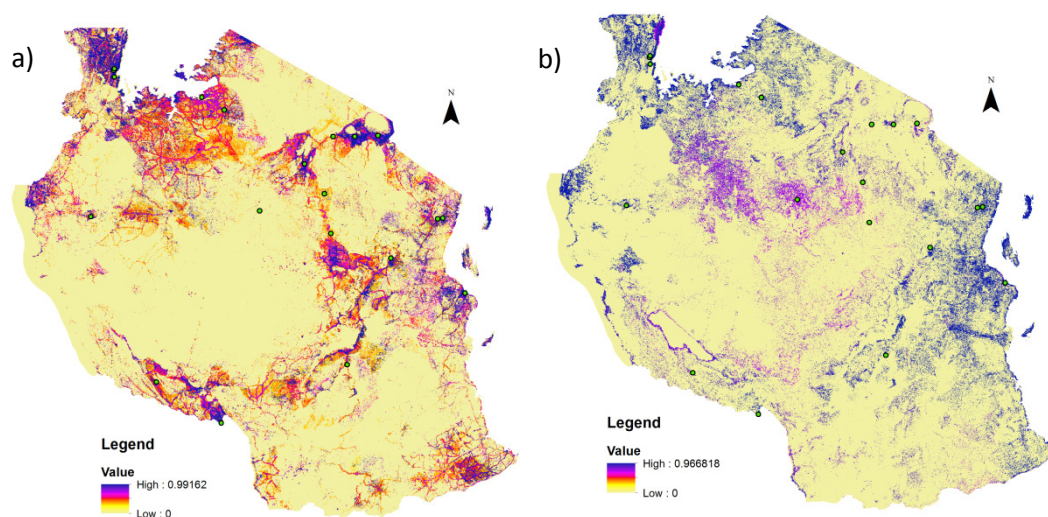**Habitat suitability maps generated**

Figure1: Habitat suitability maps for Tanzania a) for *Anopheles arabiensis* and b) for *Anopheles gambiae sensu stricto*. Occurrence points are shown in green. Human population density and land cover (specifically urban and build up, and cropland/natural vegetation mosaic land covers) contributed the most to habitat suitability for *Anopheles arabiensis*. Precipitation in the coldest quarter of the year and in the wettest month,

temperature seasonality, and elevation also contributed. Land cover (specifically croplands, urban and build up, and cropland/natural vegetation mosaic land covers) contributed the most to habitat suitability for *Anopheles gambiae ss*. Temperature annual range, precipitation seasonality and temperature seasonality also contributed.
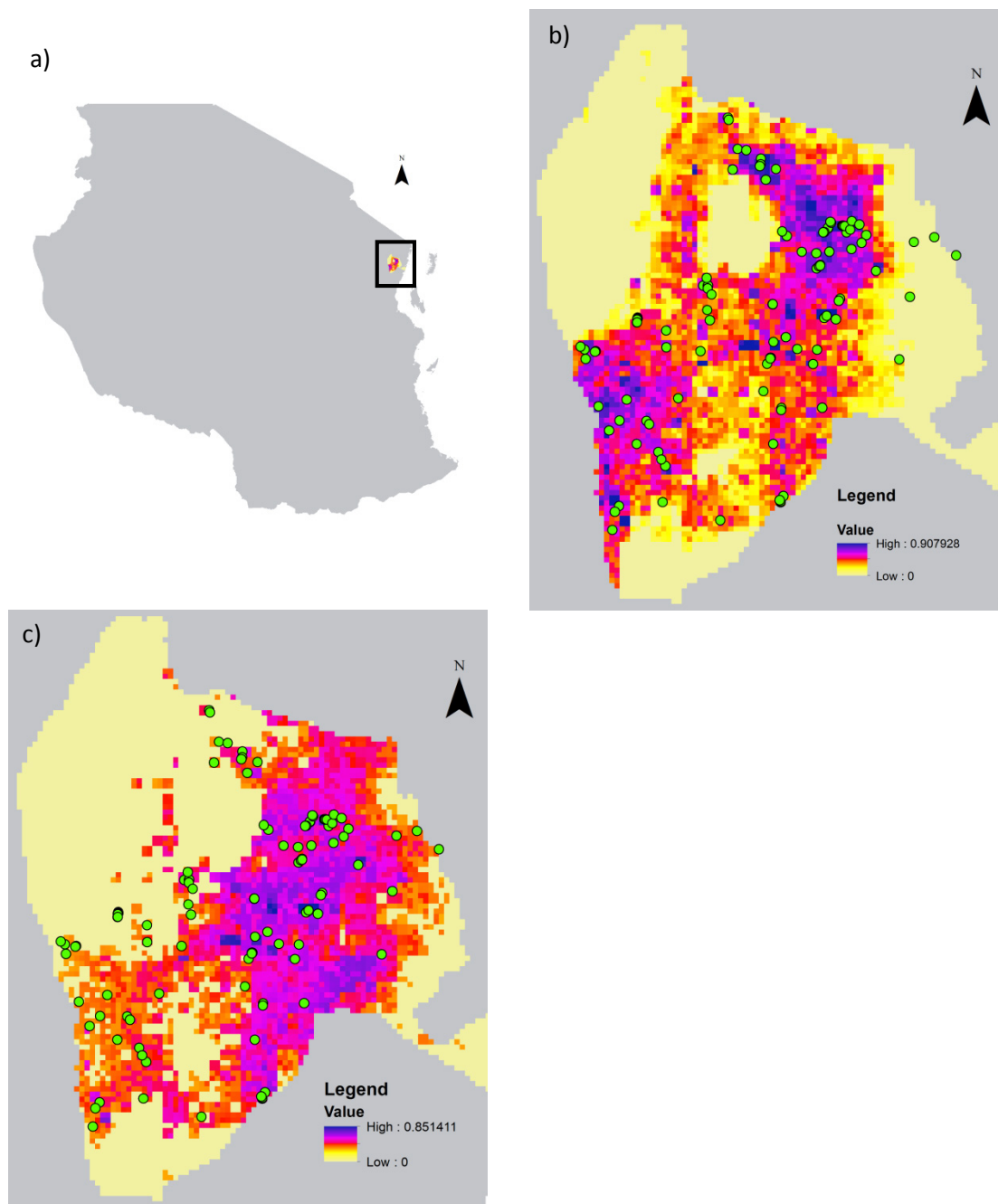

Figure2: Habitat suitability maps for Muheza district (within Tanzania – a) b) for *Anopheles sp.* and c) for *Aedes sp*. Sampling locations are shown in green. Human population density, mean diurnal temperature range, elevation and precipitation in the warmest and wettest quarters contributed the most to habitat suitability for *Anopheles sp*. Land cover (specifically savanna and cropland/natural vegetation mosaic land covers) and maximum temperature of the warmest month also contributed. Human population density, mean temperature of the wettest quarter and elevation contributed the most to habitat suitability for *Aedes sp*. Precipitation in the wettest month and land cover (specifically savanna and cropland/natural vegetation mosaic land covers) also contributed.
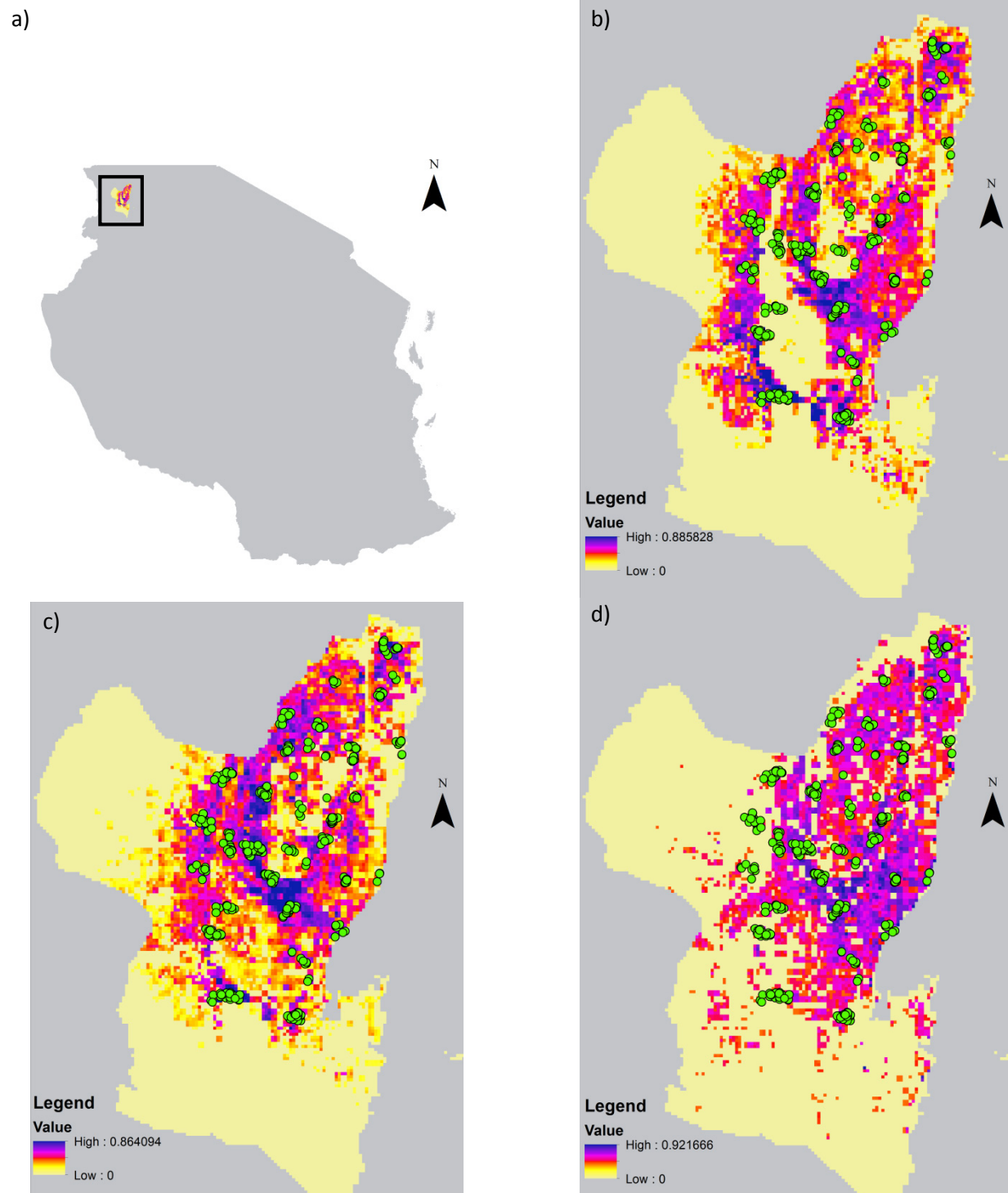
Figure3: Habitat suitability maps for Muleba district (within Tanzania – a) b) for *Anopheles arabiensis,* c) for *Anopheles gambiae ss*, and d) for *Aedes sp*. Sampling locations are shown in green. Annual precipitation and precipitation in the driest quarter contributed the most to habitat suitability for *Anopheles arabiensis*. Land cover (specifically woody savanna: negatively related to suitability), precipitation of wettest month and in the coldest quarter, temperature variability, and elevation also contributed. Annual precipitation and temperature variability contributed the most to habitat suitability for *Anopheles gambiae ss*. Precipitation in the driest quarter and land cover (specifically woody savanna: negatively related to suitability) also contributed. Human population density, precipitation in the driest quarter and mean diurnal temperature range contributed the most to habitat suitability for *Aedes sp*. Land cover (specifically croplands) also contributed.

Project implementation and management: *Briefly describe the activities supported under the project during the reporting period. Describe and discuss the research methods and analytical techniques used and any problems that arose. Indicate and explain any changes in orientation that may have occurred since the project was designed. Where applicable, comment on the financial variances which resulted from the activities of the project.*

The steps taken to achieve the five goals of the project are outlined below.

G1. A first workshop to provide advanced training in GIS and Maximum Entropy (maxent) modelling (a form of species distribution modelling to generate suitable habitat maps) to our research collaborators (Filemon Tenu – PhD candidate; Debora Kajeguka – PhD candidate; Michael Mahande – Post-doctoral fellow) and colleagues at KCMUC.

The first workshop entitled "Remote sensing modeling techniques in malaria and arbovirus control" took place September 9th to 14th, 2013. The workshop goals were to provide advanced training to research partners at KCMUC and NIMR (specifically William Kisinza – Vector Ecologist at NIMR, Filemon Tenu – PhD candidate in disease modelling at NIMR, Debora Kajeguka – PhD in arbovirus monitoring at KCMUC, Michael Mahande – post-doctoral fellow in epidemiology and biostatistics at KCMUC) in GIS, remote sensing, and Maximum Entropy (maxent) modelling. To increase the reach of the capacity building activities, the first 3.5 days were opened up to 3 additional researchers from KCMUC and NIMR (favouring women as they are currently underrepresented) Three additional researchers from KCMUC attended: Johnson Matowo (PhD student, KCMUC), Robert Kaaya (MSc student, KCMUC) and Natasha Protopopoff (PhD, Research Fellow of Entomology, London School of Hygiene and Tropical Medicine). The workshop covered acquisition of satellite imagery and global and national level GIS products(climatic grids, population data and land cover data); its processing in ArcGIS using introductory and advanced GIS functions including reprojection, map algebra, data conversion, cell size resampling, cell alignment and clipping; and use of maxent software. In addition to this, each workshop participant presented on their current research projects. The final days of the workshop were reserved for only Mr. Tenu, Dr. Mahande and Ms Kajeguka to focus specifically on modelling the habitat of *Anopheles* and *Aedes* species in the study area.

Outputs generated:
Preliminary maps of *Anopheles arabiensis* and *Anopholes gambiae s.s.* were created and analysed using data from 2006 collated by Dr. Kulkarni.

Intermediate and immediate objectives contributed to:
1a. To generate a map of suitable habitat for *Anopheles* mosquitoes (malaria vectors)

1b. To generate a map of suitable habitat for *Aedes* mosquitoes (arbovirus vectors)

3. To increase capacity of researchers to carry out high resolution vector modelling and conduct spatial statistical analyses.

Workshop objectives not achieved:
Contemporary records of occurrences of *Anopheles arabiensis*, *Anopheles gambiae s.s.* and *Aedes aegypti* were not yet acquired and brought together therefore preliminary contemporary maps for the three species could not be generated. Discussions were held to identify project sources and key collaborators from which contemporary occurrence records for *Anopheles arabiensis*, *Anopheles gambiae s.s.* and *Aedes aegypti* will be gathered. Each of the research partners at KCMUC and NIMR agreed to contact key people in their organization to assemble and share the occurrences with the group.

G4. A second workshop to provide training in backcasting to validate the species distribution models as well as spatial statistical methods to relate clinical incidences of malaria and arbovirus infections with the spatial distributions of suitable habitat for the disease specific vector.

The second workshop entitled "Modelling and statistical techniques in malaria and arbovirus control workshop" took place from October 26th to November 1st, 2013. The workshop goals were to expand upon the training provided in the first workshop on ArcGIS, data acquisition and Maxent software. The research partners in attendance were Filemon Tenu (PhD applicant in disease modelling, NIMR), Debora Kajeguka (PhD in arbovirus monitoring, KCMUC), and Michael Mahande (post-doctoral fellow in epidemiology and biostatistics, KCMUC). The first days focused on taking stock of the available data, generating preliminary *Anopheles* maps with the available data and developing a strategy for achieving the project goals. The remaining days were used to cover the topics described above. To increase the reach of the capacity building activities, the remaining days were opened up to the additional researchers from KCMUC and NIMR that participated in the first workshop. Three additional researchers from KCMUC attended: Robert Kaaya (MSc student, KCMUC) and Natasha Protopopoff (PhD, Research Fellow of Entomology, London School of Hygiene and Tropical Medicine) returned for the sessions on data extraction and spatial analysis in ArcGIS. Jovin Kitau who was unable to attend the first workshop but who already has strong GIS skills also attended these sessions. The workshop covered a refresher on the ArcGIS tools covered in the first workshop, followed by the use of tools for extracting data from GIS mapping to be used in statistical analysis, and tools for spatial analysis in ArcGIS. More specifically, we used the Hot Spot Analysis (Getis-Ord Gi*) tool for hotspot detection and the Exploratory Regression tool to examine the environmental factors related to malaria risk while accounting for spatial autocorrelation.

Outputs generated:

Contemporary occurrence records of *Anopheles arabiensis* and *Anopheles gambiae s.s.* for 2011 were assembled by Mr Tenu from the on-going resistance study conducted by NIMR. Contemporary occurrences records for *Anopheles arabiensis*, *Anopheles gambiae s.s.* and *Aedes aegypti* were made available by Dr. Protopopoff from the two-year longitudinal study in Muleba district. The geographical spread of the species occurrence records from the study in Muleba (and similar in the data that will be collected in Moshi and Muheza for Ms Kajeguka's and Mr Tenu's theses, respectively) is quite restricted relative to the geographical spread of the species occurrence records from the NIMR resistance study, which span much of the northern half of Tanzania. Consequently the decision was made to generate the species distribution maps in the following manner:

1. Species distribution maps generated:
   a) Contemporary Tanzania-wide maps for *Anopheles arabiensis* and *Anopheles gambiae s.s.* were generated using the occurrence records from the resistance monitoring study as well as one randomly selected record from each of the more detailed studies in Muleba, Moshi and Muheza.
   b) All occurrence records from the longitudinal study in Muleba were used to generate species distribution maps for *Anopheles arabiensis*, *Anopheles gambiae s.s.* and *Aedes aegypti* in Muleba district only.
   c) All occurrence records for *Anopheles arabiensis*, *Anopheles gambiae s.s.* and possibly *Aedes aegypti* from Mr Tenu's PhD field work will be used to generate species distribution maps for Muheza district only. This work has been incorporated into Mr Tenu's doctoral work.
   d) All occurrence records for *Anopheles arabiensis*, *Anopheles gambiae s.s.* and *Aedes aegypti* from Ms Kajeguka's PhD field work will be used to generate species distribution maps for Moshi district only.

2. Modelling the relationship between vector habitat and disease risk:
Modelling the relationship between vector habitat and disease risk proved to be beyond what was possible within the timeline of the grant. As part of the data collection for external projects taking place in Muleba, Muheza and Moshi, malaria prevalence is being/will be collected from households and health facilities. In the case of Moshi and Hai districts, arbovirus prevalence will also be collected from health facilities through Ms Kajeguka's doctoral work. Relating malaria prevalence and, in the case of Moshi, arbovirus prevalence to the amount of suitable vector habitat has been incorporated into the doctoral theses of Mr Tenu (for Muheza) and Ms Kajeguka (for Moshi/Hai and Muleba). This information will improve the understanding of the relationship between local environment and the likelihood of incidences of malaria relative to the incidences of arbovirus.

Intermediate and immediate objectives contributed to:
1a. To generate a map of suitable habitat for *Anopheles* mosquitoes (malaria vectors)

1b. To generate a map of suitable habitat for *Aedes* mosquitoes (arbovirus vectors)

2. To identify probable hotspots for malaria and arbovirus infections

3. To increase capacity of researchers to carry out high resolution vector modelling and conduct spatial statistical analyses.

Workshop objectives not achieved:
Backcasting of the species distribution models to assess their ability to accurately predict historic species occurrence records was not covered in the workshop. As described above, the district-level habitat maps will be validated by field work and data collected in up-coming projects.

G2. To generate preliminary habitat maps of the mosquito vectors, led by Mr. Tenu, Ms. Kajeguka and Dr. Mahande (intermediate and immediate objective 1a and 1b above).

Dr. Mahande led the development of the vector habitat maps for Tanzania, Mr. Tenu led the development of the maps for Muheza district, Ms Kajeguka led the development of the maps for Moshi and Hai districts and Dr. Desrochers led the development of the maps for Muleba district. To support Dr. Mahande, Mr Tenu and Ms Kajeguka with the development of the vector habitat maps, conference calls were held near-monthly (October 16th, 2013; December 6th, 2013; January 17th, 2014; February 25th, 2014, March 27th, 2014, April 16th, 2014). These maps can be found in the **Habitat Suitability Maps Generated** section. The regular Skype calls continued with Ms Kajeguka, Mr Kaaya, Professor Mosha and Dr. Protopopoff to support data collection, coordinate on-going collaboration and plan future projects.

G3. To conduct field work to validate the habitat maps, led by Mr.Tenu, Ms. Kajeguka and Dr. Mahande (intermediate and immediate objective 1a and 1b above).

We decided, where possible, to focus field work on targeting gaps in data availability both now and near the future to enable our partners to generate better habitat maps in their study areas. This led to most field work funds being directed toward patient and mosquito testing for arboviruses in Moshi/Hai and Muheza districts as well as increasing data collection generally in Moshi/Hai districts. It was cost prohibitive to field validate the maps for all of Tanzania and Muleba district given the size of Tanzania and the isolation of Muleba district. We obtained permission to use the data that will be collected in Muleba and Muheza in the near future by researchers at KCMUC and NIMR in order to independently validate our habitat models. Mosquitoes will be collected throughout the previous study areas and extending beyond them. However, data collection for these projects has only just started and it was not possible to conduct independent field validation of these maps in

the time frame of the current grant. However, this grant has made possible the strong research partnership between HealthBridge, the University of Ottawa, KCMUC and NIMR and thereby the continuation of this important research.

G5. A final workshop to discuss the results and plan the preparation of manuscripts presenting the results.

The final workshop took place in the second year of the project rather than the first year as expected. We decided to postpone the final workshop to discuss the results and plan the preparation of manuscripts until May, 2014 (May 20th to 28th). We made this decision for several reasons. The challenges involved in assembling and collecting species occurrence records delayed the development of vector habitat maps for the districts of Muheza and Moshi/Hai. Furthermore, our Tanzanian partners expressed a preference for support in selecting and conducting field work that took place over the last year.

The final workshop began with a refresher of the spatial statistical techniques initially covered during the second workshop in October 2013. These include tools for extracting data from GIS mapping to be used in statistical analysis, and tools for spatial analysis in ArcGIS (the Hot Spot Analysis tool for hotspot detection and the Exploratory Regression tool to examine the environmental factors related to malaria prevalence while accounting for spatial autocorrelation). We then covered the statistical methods of validating the habitat models with field data. The following days were devoted to discussion of the results as a group. This allowed us to achieve the first and second intermediate objectives of the project:
> *1. To identify environmental factors that govern the spatial prevalence of malaria relative to arboviral infections.*
> *2. To identify probable hotspots for malaria and arbovirus infections*

The workshop ended with a discussion of the manuscripts that will be prepared and their authors. We mapped the scope for three publications:
1. The environmental drivers for habitat of different vector species in Moshi and Muleba: led by Ms Kajeguka as part of her doctoral thesis
2. The strength of the habitat – malaria association in Muheza: led by Mr Tenu as part of his doctoral thesis
3. The process of species distribution modelling: strengths and limitations of using existing mosquito collections for mapping vector habitat: led by Dr. Desrochers and Dr. Kulkarni

During the workshop, we composed an abstract for the Canadian Conference on Global Health. The conference was held in Ottawa on November 2-4, 2014. Its topic was Partnerships for Global Health, a topic particularly well-suited to the nature of this project. Two of our Tanzanian partners (Ms Kajeguka and Mr Kaaya) traveled to Ottawa to attend the conference and present the findings of our project.

Project delays and forecasted expenditures: *Address other project management issues which affected the project during the reporting period and which will bear on the future of the project: forecasted expenditures for the next reporting period and until the end of the project (comment on the financial forecast included in the interim financial report, with respect to the activities planned for the next reporting period, highlighting variances from the original plan or from the last forecast);*

The delay in the timing of field work has resulted in most field data being collected in 2014 and 2015. The laboratory analyses to identify the species of mosquito and to test the mosquitoes for infection with malaria parasites or an arbovirus is currently taking place. Consequently, it was not possible to conduct independent field validation of these maps in the time frame of the current grant. However, the strong research partnership that has developed among HealthBridge, the University of Ottawa, KCMUC and NIMR as a result of this grant has made possible the continuation of this important research beyond the life of the grant.

We were able to obtain annual non-for-profit ArcGIS licenses at a substantial savings. This has left us with an additional surplus of $4,000. With your permission, we reallocated the funds for the ArcGIS software to field costs, resulting in a larger budget for data collection in the field and the ability to collect mosquitoes for more sites, which will result in better habitat suitability maps.

Robert Kaaya, a research assistant at KCMUC and new collaborator on the project, attended a symposium of the Pan-African Mosquito Control Association (PAMCA) in Nairobi, Kenya rather than the 16th International Congress on Infectious Disease in South Africa because his supervisors felt this work provide better exposure of our project. This resulted in a savings on conference costs that were used to have Robert Kaaya, in addition to Debora Kajeguka (PhD candidate also involved in the project), join us in Ottawa for the Canadian Conference on Global Health in November 2014.

For a detailed account of the funds spent, please see the attached budget summary.

Project outputs and dissemination: *Provide a listing of project outputs to date. Identify any outputs that are planned, but which have yet to materialize. Please specify what dissemination efforts were made wherever relevant. Three general categories of outputs can be identified:*

The following is a list of outputs that were developed through the grant:
Knowledge creation:
- Suitable habitat models for Tanzania and the districts of Muleba and Muheza with an analysis of environmental variables driving the distribution of vector habitat (see Figures 1 to 3).
- Information sharing and dissemination: a manuscript is being prepared presenting the results of the research portion of the project and will be submitted for peer-reviewed publication; a poster entitled "Increasing capacity in disease vector modelling to improve malaria and arbovirus mitigation strategies" was presented at the EcoHealth 2014 Conference, held from August 11-15 in Montréal (Québec, Canada); a poster entitled "Innovative approaches to solve emerging global health challenges: mapping disease vectors to target malaria and arbovirus interventions in Tanzania" was presented at the Canadian Conference for Global Health from November 2-4 in Ottawa (Ontario, Canada) by Ms Kajeguka. PDFs of both posters are attached.
Capacity-building:
- Training: 2 workshops on GIS, species distribution modelling and spatial statistics.
- Institutional reinforcement and sustainability of the research organization: New equipment: 2 laptops for the Tanzanian partners for GIS analysis and species distribution modelling.
- Increased research skills: Training guides on Introduction to GIS (and a refresher focusing on data extraction), Introduction to Species Distribution Modelling using Maximum Entropy Modelling (Maxent), and Introduction to Spatial Statistics using ArcGIS and SAM (Spatial Analysis for Macroecology). All four guides are attached.
- Training: 1 final workshop on habitat model validation and spatial statistics.

- Capacity-building of women: Ms Kajeguka (PhD candidate at KCMUC) is an integral part of all aspects of this project.

Impact: *Describe and assess any development impact that the project may have had or might be expected to have. A useful distinction can be made here between the concepts of reach and impact. Reach refers to the reception and use of the knowledge produced. Impact refers to the influence of this new knowledge on decisions or on development more generally. Special attention should be paid to the expected impact on marginalized social groups.*

The two long-term goals of this project were to increase the in-house capacity of KCMUC and NIMR for spatial analysis and vector habitat mapping, and to strengthen the budding research partnership between KCMUC (including NIMR) with the University of Ottawa and HealthBridge that began in 2009 through the development of the University of Ottawa's field course on the Ecology of East African Ecosystems by Dr. Manisha Kulkarni, which included collaboration with malaria researchers at KCMUC. We have made great strides with respect to both goals.

Our research partners (Mr Tenu, Ms Kajeguka and Dr Mahande) not only now have access to ArcGIS but have also gained indispensable knowledge on sources of useful environmental and demographic data, the variety of potential uses of GIS, how data should be collected so that it is can be used in future spatial analyses (most specifically species distribution modelling) and, most importantly, have developed a solid base in using GIS to produce maps, make spatial measurements and extract spatial data. The impact of these activities is all the greater because Mr Tenu, Ms Kajeguka and Dr. Mahande are early career researchers, meaning that they will use, improve and share these skills for years to come. In particular, both Mr Tenu and Ms Kajeguka has incorporated vector modelling and spatial analyses into their doctoral theses. By extending the training sessions to include other researchers at KCMUC, we have also introduced a several other researchers at KCMUC and NIMR to the value of GIS, spatial analysis and habitat modelling as well as deepening the skills of a number of researchers that had already been introduced to GIS. We now also collaborate extensively with Robert Kaaya, who attended and benefited from all three workshops. He is a recent MSc graduate and has been indispensable in carrying out the mosquito collections in the field. We hope he will agree to pursue doctoral studies on this topic and will continue to collaborate with HealthBridge and the University of Ottawa.

The partnership between KCMUC (including NIMR) and the University of Ottawa and HealthBridge has developed into a very productive collaboration. As a result of the research and collaboration funded our Canadian Partnerships grant, we received a Grand Challenges Canada, Stars in Global Health Phase 1 grant to expand the project and develop seasonal models of malaria and arbovirus risks. We will integrate these model predictions of the relative risk of malaria and arbovirus infections into a diagnostic Smartphone application for diagnosis of febrile illness (fever). We are currently collaborating with Dr. Protopopoff to further capitalize of the data collected during the 2 year study (2011-2012) longitudinal study in Muleba, where mosquitoes were captured monthly from households in 40 villages throughout the district in 2011(Protopopoff *et al*, Malaria Journal, 2013). We are using these data to generate maps of suitable habitat for *An. arabiensis*, *An. gambiae ss* and *Aedes sp*. on a monthly basis at a higher spatial resolution (individual pixels will represent 30m on the ground rather than 250m). We are also in regular contact with Professor Mosha, Dr Kisinza, Dr. Protopopoff, Ms Kajeguka, Mr Tenu and Mr Kaaya for proposal development to continue both the field data collection and the vector modelling in larger areas.

Recommendations: *Include in this section a summary of any recommendations that you would like to make to IDRC with respect to the administration of the project, related to the scope, duration, or budget. These recommendations derive from the background information provided under Project implementation and management.*

Reviewers of our initial grant application expressed concern regarding the large proportion of the budget that was allocated to travel to Tanzania by the Canadian researchers involved in the project (Dr. Desrochers, Dr. Kulkarni and Dr. Kerr). The reviewers made the sound point that all three of us were not required for each trip

and also made the wise suggestion to increase the number of researchers in Tanzania that would benefit from the training. We have observed that in-person meetings are far more productive than conference calls and the remote collaboration, more generally. For future projects, it may be most beneficial for fewer researchers to travel (as suggested by the reviewers) but to travel more frequently, though we acknowledge the negative environmental and fiscal impact this would have.