Developing risk assessment maps for Schistosoma haematobium in Kenya based on climate grids and remotely sensed data

Kelsey Lee McNally

*Louisiana State University and Agricultural and Mechanical College*

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the Veterinary Pathology and Pathobiology Commons

**Recommended Citation**


https://digitalcommons.lsu.edu/gradschool_theses/1851

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
Acknowledgements

I would like to thank Dr. Jack Malone for working with me as my major professor. He has been instrumental in providing me with any data and telling me I was doing something wrong before it was too late. Next I would like to thank Dr. James Miller and Dr. Oscar Huh for being on my committee. Thank you to Jennifer McCarroll and DeWitt Braud for technical assistance and problem solving throughout my project. Thank you to Pat Smith and Seklau Wiles for their encouragement me to explore my natural potential. Thank you to Emanuele Brianti and Prixia del Mar Nieto for making me laugh and relax when I was stressed out and everything seemed to be going wrong. Thank you to Michael Kearny who was a great help with the statistics. Finally I want to thank my mother Molly McNally and my sister Lacey McNally for everything they have done for me and always willing to do whatever I asked of them.
# Table of Contents

Acknowledgements........................................................................................................... ii

List of Tables....................................................................................................................... iv

List of Figures................................................................................................................... v

Abstract............................................................................................................................. vii

Introduction....................................................................................................................... 1

Methods and Materials.................................................................................................... 10

Results............................................................................................................................... 17

Discussion......................................................................................................................... 30

Conclusion and Significance............................................................................................ 37

References......................................................................................................................... 38

Vita.................................................................................................................................... 42
List of Tables

1. Multiple-regression analysis output for monthly averages of climate and calculated factors…………………………………………………..24

2. Multiple-regression analysis output for annual totals of Climatic and calculated factors…………………………………………………..24
List of Figures

1. Data layers that were used to construct the environmental risk assessment models ................................................................. 10

2. Relationships of mortality rate (d), and effective recruitment rate (b), to mean water temperature ............................................. 11

3. Stratified growth curve showing the rate of development compared to temperature for S. haematobium ..................................... 12

4. Black areas have a monthly mean temperature above 28.5°C, which causes gonadal atrophy .................................................... 13

5. Potential generations of S. haematobium per year (Cumulative Annual GDD/268) ................................................................. 18

6. Risk index predicted distribution map, after correcting for the effect of increasing levels of heat stress ................................... 19

7. Four charts showing the value ranges of Tmax and NDVI for the wet season and the dry season .................................................. 20

8. Areas within the Tmax ‘suitable’ range of 15°C to 29°C ................................................. 21

9. Areas of adequate moisture based on NDVI map queries ................................................. 21

10. Area determined to be potential habitat for snail development based on meeting both NDVI and Tmax .................................. 22

11. Areas selected by Desktop GARP as preferred habitat ................................................. 23

12. Risk assessment maps that combine suitable habitat as determined by using AVHRR sensor data and the growing degree day model .................................................. 25

13. Risk assessment maps that combine suitable habitat as determined by using the AVHRR sensor data and the modified risk index .................................................. 26

14. GDD model that did not use the PPE>0.5 threshold ................................................. 28

15. The modified heat stress model, without the PPE>0.5 threshold ................................................. 29
16. Linear relationship of temperature to the rate of 
   S.mansoni development .................................................................31

17. S.haematobium prevalence compared to annual 
   average Tmax for East Africa ..........................................................32
Abstract

It is important to be able to predict the potential spread of water borne diseases when building dams or redirecting rivers. This study was designed to test whether the use of a growing degree day (GDD) climate model and remotely sensed data (RS) within a geographic information system (GIS), could be used to predict both the distribution and severity of *Schistosoma haematobium*. Growing degree days are defined as the number of degrees centigrade over the minimum temperature required for development. The base temperature and the number of GDD required to complete one generation varies for each species. A monthly climate surface grid containing the high and low temperature, rainfall, potential evapotranspiration (PET), and the ratio of rain to PET was used to calculate the total number of GDD provisional on suitable moisture content in the soil. The latitude and longitude for known snail locations were used to create a point file. A 5km buffer was made around each point. Mean values were extracted from buffer areas for Advanced Very High Resolution Radiometer (AVHRR) data on maximum land surface temperature (Tmax) and normalized difference vegetation index (NDVI). The values for Tmax ranged from 15-28 and the NDVI values were 130-157. A map query found all areas that meet both criteria and produced a model surface showing the potential distribution of the vectors for this disease. Results indicate that the GDD and AVHRR models can be used together to define both the distribution range and relative risk of *S. haematobium* in anticipated water development projects and for control program planning and better allocation of health resources in endemic vs. non-endemic areas.
Introduction

Schistosomiasis is one of the World Health Organization’s (WHO) great neglected diseases. It is the second most prevalent tropical disease, with 200 million people infected and 500 million at risk (Lengeler et al., 2002). The great majority (80-85%) of schistosomiasis is found in sub-Saharan Africa (Bergquist, 2002). It is important to create some method of separating areas of high risk from areas of low risk for control programs. Schistosomiasis is spread by water contact of human hosts in freshwater habitats of snails suitable for propagation and transmission of the parasite.

*Schistosoma haematobium* has an indirect life cycle that requires aquatic snail intermediate hosts of several Bulinus spp. Suitable snail hosts are penetrated by miracidia that hatch in fresh water aquatic environments from eggs shed in the urine of infected humans. The miracidia have 8-12 hours to infect a suitable Bulinus spp. snail host. In the snail, the miracidium then develops into sporocysts and then to cercariae. The sporocyst can produce up to 600 daughter sporocysts. The daughter sporocysts migrate to the digestive gland of the snail host to begin producing the cercariae stage. The snail can shed approximately 100’s of cercariae/day up to 18 days (http://martin.parasitology.mcgill.ca/jimspage/biol/schisto.htm). Temperature effects the amount of time needed to complete this portion of the life cycle. Cooler temperatures decrease the rate of development while warmer temperatures up to the optimum temperature 22°C. Pflüger (1981) describes this portion of the schistosome life cycle as being the most restrictive with respect to temperature. The free swimming cercariae released from snail hosts infect the human host by penetrating the skin. The cercariae shed their bifurcated tail, and the resulting schistosomula are carried through the blood
stream to the lungs for approximately 8 days. The schistosomula are then transported to the liver where they mature to adults and begin pairing as adult. The paired adults begin producing eggs that pass into adjacent tissues, while many eggs penetrate through the mucosa of bladder. The eggs of *S. haematobium* are shed in the urine (Ross, 2002). Eggs contaminate fresh water due to the lack of a proper sewerage system or defecating close to a water body. The eggs, once in water, then hatch because of the osmotic pressure releasing the miracidia to the water to infect more snail hosts.

Advances in technology have allowed geographical information systems (GIS) and remote sensing (RS) to be used in the epidemiology of disease to create risk assessment maps. John Snow first used GIS in 1854, in a public health application of an outbreak of cholera in London (Loslier, 1995). Because of the lack of spatial data sources and the limited capability of computer software, GIS has not been used at its full potential, until recently. Technological advances in computers, software, and data availability have led to development of a number of GIS medical applications that allow evaluation of both spatial and temporal relationships of the environment and disease agents. In the last decade, GIS has been used to create risk models for a number of environmentally sensitive diseases including malaria, onchocerciasis, rift valley fever, fascioliasis, and African trypanosomiasis (Brooker, et al. 2001, Malone et al., 1995, 2003). Environmental factors such as climate, satellite sensor data, elevation, land use, soil type, and other map data are overlaid on a base map of standard geographic projection and scale. A series of points can then be added and used to extract data for all layers of environmental data and interactively analyzed by GIS query and statistical methods. These points can be cities,
weather stations, or locations where health data was collected in previous or current field research.

Remote sensing data from earth observing satellites or aircraft can be used to gather ‘surrogate’ information on climate, vegetation, soil moisture, and other environmental features. Remote sensing has not been widely used in public health, until recently due to the coarse resolution of early sensors and the cost of obtaining satellite images. Environmental satellites have been launched into space for 40 years, but were originally limited to use for the military and weather observations (Huh and Malone, 2001). Among the advantages of using RS are that data can be provided on areas that are not assessable, it offers the possibility of global coverage, and data can be collected, processed, and used in near real time. Weather stations can only provide data for small areas, and therefore it can require hundreds or thousands of climate stations to effectively cover a given country in the same detail as satellite imagery.

Currently, the Advanced Very High Resolution Radiometer (AVHRR) sensor is one of the most widely used remote sensing systems used in parasitological and epidemiological studies. This system uses 5 bands with bands 1 and 2 being in the visible portion of the energy spectrum, bands 4 and 5 being thermal infrared data, and band 3 being mid-range infrared data, all with a pixel resolution of 1.1km² (Huh and Malone, 2001). The first AVHRR sensor satellite was launched in 1978, providing a long-term archive of satellite data that can be used to map the distribution and abundance of diseases and vectors and to document the spread of a given disease. A limitation of using AVHRR is the 1.1km² resolution and the fact that snails and the parasites they carry often live in water bodies and streams smaller than 1.1km².
The effects of environmental factors on the development of the intermediate hosts of schistosomiasis were described by Malek in 1958. A suitable range was defined for each factor and the limiting effect it has on the snail. It is reported that *Bulinus globosus* and *Bulinus physopsis* can survive 12 months of drought in the laboratory and 45 days in its natural habitat (Malek, 1958).

Many studies have been done to evaluate and predict the effects of irrigation projects on snails and schistosomes with mixed success. For example in 1976, the United States Agency for International Development (USAID) did an assessment on the risk of schistosomiasis before a water development scheme began in the Senegal River Basin and predicted the river basin was at low risk for the development of this disease. Two dams were built, one in 1985 and a second in 1989, which changed the environment in the Senegal River Basin and led to one of the worst outbreaks of schistosomiasis ever recorded (Southgate, 1997). This was likely to occur, since data cannot be recorded for every location or every point in time (Kitron, 2000) and thus field assessments were incomplete. This event and earlier experience with other water development projects that ultimately led to increased risk of schistosomiasis, notably the Aswan Dam in Egypt and the Akosombo Dam in Ghana (Hunter et al., 1982), indicate the need for more accurate risk assessment methods.

The construction of the Aswan Dam altered the environment, creating unfavorable conditions for *Bulinus* sp. and favorable conditions for *Biomphalaria* spp. Under pre-dam conditions of a single seasonal irrigation period that were more suitable for the development of *Bulinus* sp., the prevalence in the Nile Delta of *S. haematobium* was 74% in 1935 and the prevalence of *S.mansoni* was 3.2% (Abdel-Wahab, 1979). In
1979, when the human population was retested, the prevalence relationship had reversed; *S. haematobium* prevalence was 2.2% and *S. mansoni* prevalence was 73%.

Irrigation and water development is thus known to be highly influential in the spread and increased relative severity of schistosomiasis. Sturrock (1965) monitored nine irrigations schemes in Tanzania to determine the effect of these water developments on the establishment of snail populations. Each scheme used different methods of water control and maintenance leading to the establishment of different assemblage of snail species in each scheme. In 1977, established irrigation schemes in Ethiopia were surveyed to measure the prevalence of schistosomiasis, the effects of water temperature, and elevation on the snail population. The results of this study showed that *Biomphalaria pfeifferi* were found in large numbers, while only a few snails from the *Bulinus* spp were found in the canals (Kloos and Lemma, 1977). It was reported that resettlement was another method of introducing infected people to uninfected areas. In many cases the schistosomes died and did not create new endemic areas (Kloos, 1990).

The ecology of *S.mansoni* and *S. haematobium* was mapped for Ethiopia by Kloos (1988) to show the relationships between prevalence and environmental factors, both natural and man-made. An elevational limit of 800m-2200m above sea level was established as suitable elevation for *S.mansoni* due to hot water temperatures below 800m and cold water temperature above 2200m (Kloos et al., 1988). *S. haematobium* was found below 800m because the *Bulinus* spp. in Ethiopia favor warm water temperatures. *S.mansoni* prevalence declines when the elevation is above 2000m (Kloos et al., 1988). In addition to describing the elevational requirements for the intermediate hosts, Kloos also identified the specific species capable of transmission for various region in Ethiopia.
The effects of temperature on prevalence and development of schistosomes have been studied in the laboratory by a number of authors. Anderson et al. (1979) found that the rate of development increased with temperature as a linear relationship. The suitable temperature range for larval development in the laboratory was between 15°C to 35°C, with an optimum temperature for intramolluscan development of 25°C. The study also determined that infection considerably reduced the life span of the snail. Woolhouse et al. (1990) describes the fecundity rate as being dependent on temperature by modeling a bell curve relationship with the optimum temperature peak at 20.6°C.

The thermal development requirements of \textit{S. haematobium} and \textit{S.mansoni} have been reported in detailed studies by Pflüger (1981). A minimum temperature and maximum temperature for development of the parasite was established as well as the number of GDD required to complete one life cycle in the snail host for both \textit{S.mansoni} and \textit{S. haematobium}. The rate of development, as influenced by temperature, was shown to have a linear relationship starting with the minimum temperature (15.3°C for \textit{S. haematobium}) to an optimum temperature (22°C for \textit{S. haematobium}). Again, it was stated that there is an inverse relationship of development and temperature above the optimum temperature. Parasitic development within the snail has the most restricted range of temperatures throughout the life cycle (Pflüger, 1981). The effects of temperature on \textit{Bulinus globosus} in field studies in the Kenya coastal areas were reported by O’Keeffe, (1985); snail populations were limited when temperatures rise above 28.5°C, an effect that was attributed to gonadal atrophy. The optimum temperature for population growth was shown to be 25°C. Rainfall and water temperature were reported
to be the most important variables for controlling the population of *B. globosus* (O’Keeffe, 1985).

The use of GIS and RS for the prediction of schistosomiasis was first attempted in the Philippines and the Caribbean by Cross et al. (1984). Data collected by weather stations, topographic information, Landsat data, and disease was combined to create a model. This model had 93.2% correct classification of known endemic areas in the Philippines. The model was not accurate when classifying the environment for schistosomiasis in the Caribbean (Cross et al., 1984).

GIS and remote sensing (RS) were later used to evaluate risk of schistosomiasis in the Nile River delta of Egypt (Malone, 1994) for use in control program management for schistosomiasis. A regional scale model was created for the Nile Delta on the potential distribution and abundance of *S. mansoni* by Malone et al. (2001) who used AVHRR to create diurnal temperature difference maps which indicated thermal-hydrological regimes that favored the snail host and transmission of the parasite. *Bulinus* spp. was suggested to be more tolerant of drought and high temperatures than *Biomphalaria* spp, the snail host for *S. mansoni* (Malone, 1994, 1995).

The habitats for *Bulinus globosus* and *Biomphalaria pfeifferi* were described in Zimbabwe using Tmax and NDVI from the AVHRR sensor. The seasonal distribution was monitored using RS to determine which months had the most influence on transmission. The optimum range of NDVI was 128-160 for identifying areas of high prevalence (= 5%) (Mukaratirwa, 1999).

Risk assessment maps were created in 2001 for *S. haematobium* in Tanzania by use of RS and logistic regression, regions were identified as being risk (Brooker, 2001).
Remote sensing was used to define ecologic zones and then produce risk assessment maps for each zone. These models were then validated by using a questionnaire to determine prevalence in a given area (Brooker, 2002). A risk assessment model has been developed for *S.mansoni*, in Bahia, Brazil, using GIS. The latter model evaluated several agro-climate and environmental factors to find areas that met the life cycle requirements of *S.mansoni* (Bavia, 1999).

Geospatial studies on other diseases suggest new approaches for the use of GIS for better decision making toward schistosomiasis control. The spatial distribution of filariasis in the Nile Delta has been described using RS to identify environmental requirements for this disease. Normalized difference vegetation index (NDVI) combined with temperature were used to determine suitable mosquito habitat areas (Crombie, 1999, Thompson et al.,1996). These studies used AVHRR and the Landsat sensor systems, respectively, to measure soil moisture parameters that determine potential breeding habitat.

For malaria, an environmental risk assessment model was recently used for prediction of malaria risk for Eritrea in 2002. A growing degree day model was used to determine the potential severity of malaria which was divided into a highland pattern and a coastal transmission pattern based on temperature and water balance parameters using a 5km² climate grid (Malone et al., 2003). A similar approach was successfully used to determine the potential severity of *S.mansoni* in Kenya (Malone and Corbett, 2002). The influences of resettlement were described for Malaria, yellow fever, onchocerciasis , and trypanosomiasis in Ethiopia. An upper limit for elevation was defined as 2000m for many tropical diseases in Ethiopia (Kloos, 1990).
The impetus for the current study on *S. haematobium* in Kenya arose from results of earlier work described above and the availability of GIS-compatible resource datasets from two sources: 1) A ‘minimum medical GIS database’ for the IGAD/Nile region of East Africa, including Kenya (Malone et al., 2001a,b), and 2) a database from a very complete recent survey of the freshwater snails of southern Kenya (Loker et al., 1993).

The Global Network on Schistosomiasis Information Systems and Control of Snail-Borne Disease (www.GnosisGIS.org) has fostered use of satellite climatology in predictive models for schistosomiasis and fascioliasis. One of the purposes of this network is to create and maintain a “Minimum Medical Database” (MMDb), a compilation of environmental and parasitological data needed to develop risk assessment models for a variety of diseases and parasites. The MMDb includes AVHRR data, on NDVI and land surface temperature (Tmax), along with other data components such as a Food and Agriculture Organization (FAO) soil database, topography, and snail distribution for East Africa (Bergquist, NR, 2002).

Loker et al. (1993) performed a systematic survey in southern Kenya to map the locations of fresh water snails and determine the preferred habitat. Snails were collected and identified over a three year time period in natural and man-made water bodies. Elevation and habitat type were recorded as well as the number of each species found. The latitude and longitude were recorded so that these points could be used in a geographic information system (GIS). The MMDb database, supplemented with geospatial data from other sources, and the Loker survey database were used in the current study to initiate development of a risk assessment model for Kenya using GIS/RS methodologies.
Methods and Materials

Two approaches were taken to develop separate GIS model components which were later combined to produce the final model: (1) a climate-growing degree day model and (2) a remote sensing model using earth observing satellites. The databases included in the GIS constructed for use in these studies is shown on Fig.1.

![Diagram showing regional risk models with data layers](image)

**Fig.1.** Data layers that were used to construct the environmental risk assessment models for *S. haematobium*.

Growing Degree Day Model

A growing degree day (GDD) model was developed using a MMDb climate surface 5km² grid that had originally been obtained from the Almanac Characterization Tool 3.0 (ACT 3.0) (Corbett et al., 2001), a comprehensive database on climate, demography, topography, infrastructure, hydrology, and other environmental parameters for East Africa.
Growing degree days were calculated as the number of degrees above 15.3°C, the minimum temperature required for development of *S. haematobium* (Pflüger, 1981) times the number of days per months ((mean temperature-15.3)*days of the month). Growing degree days were accumulated only under conditions in which a soil moisture water balance was above a threshold considered to be needed to allow the snail and the parasite to survive in a given environment. A precipitation:potential evapotranspiration ratio (PPE) greater than 0.5 was set as the criteria to indicate the presence of an adequate amount of surface water or soil moisture based on the water balance in the top 25cm of soil water holding capacity. A total of 298 GDD is required to complete one *S. haematobium* development cycle in the snail host. GDD/298 represents the number of potential generations per year that could occur for each grid cell.

**Modified Risk Index Model**

It is known that development rates of the free-living stages of many parasites increase with temperature until the optimum temperature is reached, then decrease at temperatures that exceed the optimum due to heat stress (Fig.2), although this factor has been difficult to measure in the field (Andrewartha and Burch, 1954, Woolhouse, 1990).

Fig. 2 Relationships of mortality rate (d), and effective recruitment rate (b), to mean water temperature for *B.globosus* based on a mathematical model (Woolhouse, 1990).
The effect of heat stress was estimated by stratifying the growth curve into 7 classes assuming the population development-survival balance will lead to population growth at a given rate on either side of the optimum temperature (Fig.3). The water balance threshold of PPE>0.5 was used to indicate areas with adequate moisture for life cycle progression in each month. Monthly values were then summed to yield a cumulative risk index value.

![Stratified growth curve showing the rate of development compared to temperature for S. haematobium.](image)

**Fig.3.** Stratified growth curve showing the rate of development compared to temperature for *S. haematobium*.

**Masking Unsuitable Areas**

A temperature mask was created based on the incompatibility of high temperatures with snail survival over 28.5°C temperature (O’Keeffe, 1985) by GIS query analysis of the temperature grid to identify grid cells with a mean temperature above 28.5°C. Since the mean temperature varies seasonally, the query was done on a monthly basis. Four months are displayed in Figure 4 to represent seasonal pattern change.
Fig. 4. Black areas have a monthly mean temperature above 28.5°C, which causes gonadal atrophy in *B. globosus* (O’Keeffe, 1985). (a) January (b) March (c) May and (d) September.
AVHRR Remote Sensing Models

Advanced Very High Resolution Radiometer (AVHRR) data was obtained (Malone et al., 2001) for 1992-1993 and 1995-1996 via the Internet from the United States Geological Survey (USGS) global 1km² website (http://edcdaac.usgs.gov/1KM/1kmhomepage.html) for daytime land surface temperature (Tmax) and Normalized Difference Vegetation Index (NDVI). NDVI is an index that ranges from -1 to 1, that was rescaled for use by the global 1km² website to yield NDVI values of 0-200. The primary purpose of NDVI is to measure vegetation health, however, NDVI can also be used to estimate soil moisture based on vegetation health (Crombie, 1999, Huh and Malone, 2001). NDVI based on data from the AVHRR sensor is calculated by using channels 1 and 2 ((channel 2- channel1)/(channel 2+ channel 1)) (Huh and Malone, 2001).

The AVHRR data was downloaded at dekadal intervals (every 10 days) that had already been processed by the USGS to eliminate cloud cover using an algorithm that records the highest pixel value for each pixel during the 10-day period. All satellite images were calibrated and georeferenced to a geographic decimal degree latitude and longitude coordinate system using ERDAS Imagine 8.6 image processing software. The dekadal data were then combined to create monthly composite maps by averaging the images together:

\[((Image1+Image2+Image3)/3)=\text{monthly composite}\]

Monthly data were averaged together to create wet season (Oct-March), and dry season (April-Sept) composite maps, and an annual composite map. Composite images were incorporated into an ArcView 3.3 GIS project and converted to a grid file for further
analysis using a grid cell size of 1.1km², the same spatial resolution as the original AVHRR data.

Buffer zones with a 5km diameter were created in ArcView 3.3 GIS centered on all survey points for *B.globosus*. A separate set of buffers was centered on survey points containing *B.physopsis*. These buffers were used to extract mean values for both Tmax and NDVI. Since each pixel in the AVHRR image represents 1.1km², ArcView 3.3 GIS calculated the mean value for both NDVI and Tmax for the entire buffer area. Based on the mean value range, a map query was then done to select grid cells of the composite maps consistent with the range of known infected sites. The intent was to show areas where *B.globosus* and *B.physopsis* can occur based on the value range of Tmax and NDVI at known endemic sites.

**Ecological Niche Modeling**

Using the Genetic Algorithm for Rule-Set Prediction (GARP), an ecological niche for *S haematobium* was predicted based on environmental factors, including high temperature, low temperature, rainfall, evapotranspiration, topography, and AVHRR data. The snail survey dataset of Loker et al. (1993) was used as the positive point location records for this program. GARP uses 50% of the point locations for training and the remaining 50% as a test dataset, using either 1000 iterations or until an operator specified convergence level is reached in the rule-selection process. The predictive value is based on 1250 points taken from test data and 1250 points randomly selected from within the study area. Logistic regression is performed after each iteration to determine if the rule should be incorporated into the model (Peterson, 2002, 2003). An ArcView grid
is created based on the final model used by GARP. This grid can then be input into a GIS and compared to existing models.
Results

GDD and Modified Risk Index

The range of potential generations per year throughout Kenya was -8 to +10 when a water balance of PPE>0.5 was considered to be a limiting factor (Fig. 5). The area surrounding Lake Victoria and along the southern coast had 6-10 generations per year. The eastern border, the area around Lake Turkana, and the highlands had values less than 2 generations per year. For the remainder of the country 2-6 potential generations of *S. haematobium* were predicted developing in the snail hosts per year.

For the modified risk index model that accounted for the effect of heat stress, risk index value ranges of 0-15 (very low), 16-30 (low), 31-46 (moderate), 47-61 (high), and 62-77 (very high) (Fig.6) were predicted, conditional on water balance values of PPE>0.5. The area surrounding Lake Victoria is the only region with both high and very high risk by this model. The southern part of Kenya along the coast, and regions in the highlands were predicted to have moderate risk. The remainder of the southern region of Kenya had low risk while the northern regions had very low risk.

AVHRR Remote Sensing Model

The mean values for Tmax within the 5km buffers centered on snail survey data points ranged from 6.5-28.5°C for the wet season, and 7.5-29°C for the dry season based on 1992/1995 AVHRR composite seasonal maps (Fig.7). The range used for the map queries was 15?C-29°C. The lower limit of this temperature range was chosen because of the life cycle demands of the schistosome. The upper limit was used based on the fact that no snails were found at higher temperatures and also because of reported gonadal atrophy above 28.5°C (O’Keeffe, 1985).
Fig. 5. Potential generations of *S. haematobium* per year (Cumulative Annual GDD/268) that can occur based on monthly mean temperature of > 15.3°C, conditional on a water balance ratio (Rain/Potential Evapotranspiration) of >0.5.
Fig. 6. Risk index predicted distribution map, after correcting for the effect of increasing levels of heat stress (>optimum temperature) on the model output of the GDD-based generations per year model for *S. haematobium*, conditional on water balance values of PPE>0.5.
Fig. 7. Four charts showing the value ranges of $T_{\text{max}}$ and NDVI for the wet season and the dry season (a) $T_{\text{max}}$ values for known *Bulinus* spp. locations during the wet season (b) $T_{\text{max}}$ values for known *Bulinus* spp. locations during the dry season (c) NDVI values for known *Bulinus* spp. locations during the wet season (d) NDVI values for known *Bulinus* spp. locations during the dry season. All *Bulinus* spp. locations were based on surveys performed by Loker et al. (1993).

The map queries indicated almost the entire country fell within the suitable thermal range 15°C-29°C (Fig. 8). The NDVI mean values extracted for the 5km buffers were 35-160 for the wet season, and 56-157 for the dry season based on the 1992/1995 composite seasonal map data (Fig. 7). Values of NDVI used for final map queries were 125-160 for the wet season, and 130-157 for the dry season (Fig. 9). Buffer query areas with low mean NDVI values (<100) were eliminated from consideration if part of the buffer zone being averaged included a water body of nil NDVI value. Large water bodies register as 0 for NDVI due to the fact that there is little or no vegetation in the middle of a body of water. Another map query, using boolean logic, shows areas that met both criteria (NDVI and $T_{\text{max}}$), thus being suitable habitat for a snail population to survive (Fig. 10).
Fig. 8. Areas within the Tmax ‘suitable’ range of 15°C to 29°C for (a) wet season and (b) dry season as determined by dekadal AVHRR satellite sensor data.

Fig. 9. Areas of adequate moisture based on NDVI map queries (a) wet season with a range of 125-160 (b) dry season with a range of 130-157.
Fig. 10. Area determined to be potential habitat for snail development based on meeting both NDVI and Tmax suitable range criteria for the (a) wet season and (b) dry season.

Ecological Niche Modeling

Two seasonal ecological niche models for the distribution of *S. haematobium* were derived based on temperature, rainfall, topography and remotely sensed data from the AVHRR sensor (Fig. 11). Convergence at a threshold of 0.05 was reached after 20 iterations for the wet season, and 25 iterations for the dry season. Once convergence was reached, Desktop GARP performs a Chi square analysis for a “best fit” ecological model for each season. The p values from the Chi square analysis for each model were 9.13e -18 for the wet season, and 3.95e-15 for the dry season.
Fig. 11. Areas selected by Desktop GARP as preferred habitat based on climatic factors, topography, and the locations of snails (Loker et al, 1992) for the (a) wet season and (b) dry season.

Statistical Analysis

Multiple-regression analysis was done to determine which environmental and calculated factors had the highest impact on the distribution and abundance of the *Bulinus* spp. population. This was done for both a monthly average (Table 1) and annual totals (Table 2) of these factors. In both cases, the growing degree day model (WETGDD, $p=0.0029, 0.0032$) and the modified heat stress model (WETRANK, $p=0.0011, 0.0014$), as measured by the PPE$>0.5$ threshold, were the most significant factors.
Table 1. Multiple-regression analysis output for monthly averages of climatic and calculated factors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Error</th>
<th>Type II F value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-26.96707</td>
<td>57.50568</td>
<td>193.28656</td>
<td>0.22</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.01888</td>
<td>0.04634</td>
<td>145.85476</td>
<td>0.17</td>
</tr>
<tr>
<td>GDD w/o PPE&gt;0.5</td>
<td>-0.66079</td>
<td>0.41591</td>
<td>2218.6384</td>
<td>2.52</td>
</tr>
<tr>
<td>Low temperature</td>
<td>-0.71718</td>
<td>6.20303</td>
<td>11.74915</td>
<td>0.01</td>
</tr>
<tr>
<td>PET</td>
<td>-2.15153</td>
<td>1.03628</td>
<td>3788.7754</td>
<td>4.31</td>
</tr>
<tr>
<td>PPE</td>
<td>-211.911</td>
<td>162.1779</td>
<td>1500.6589</td>
<td>1.71</td>
</tr>
<tr>
<td>Heat stress w/o PPE&gt;0.5</td>
<td>-4.65159</td>
<td>9.30496</td>
<td>2196.4861</td>
<td>0.25</td>
</tr>
<tr>
<td>GDD w/ PPE&gt;0.5</td>
<td>0.63705</td>
<td>0.20967</td>
<td>8113.5196</td>
<td>9.23</td>
</tr>
<tr>
<td>Heat stress w/ PPE&gt;0.5</td>
<td>-19.72045</td>
<td>5.83661</td>
<td>10034.1104</td>
<td>11.42</td>
</tr>
<tr>
<td>High temperature</td>
<td>19.32259</td>
<td>7.70205</td>
<td>5531.8949</td>
<td>6.29</td>
</tr>
</tbody>
</table>

Table 2. Multiple-regression analysis output for annual totals of climatic and calculated factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Error</th>
<th>Type II F value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-127.807</td>
<td>181.90179</td>
<td>444.36276</td>
<td>0.49</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.0379</td>
<td>0.0551</td>
<td>425.9589</td>
<td>0.47</td>
</tr>
<tr>
<td>GDD w/o PPE&gt;0.5</td>
<td>-0.05874</td>
<td>0.03602</td>
<td>2393.398</td>
<td>2.66</td>
</tr>
<tr>
<td>Low temperature</td>
<td>0.17846</td>
<td>0.66247</td>
<td>6531983</td>
<td>0.07</td>
</tr>
<tr>
<td>PET</td>
<td>-0.16213</td>
<td>0.09338</td>
<td>2713.172</td>
<td>3.01</td>
</tr>
<tr>
<td>PPE</td>
<td>-14.93816</td>
<td>15.5069</td>
<td>835.29665</td>
<td>0.93</td>
</tr>
<tr>
<td>PRE</td>
<td>0.05938</td>
<td>0.14271</td>
<td>155.86371</td>
<td>0.17</td>
</tr>
<tr>
<td>Heat stress w/o PPE&gt;0.5</td>
<td>-0.42079</td>
<td>0.78435</td>
<td>259.05802</td>
<td>0.29</td>
</tr>
<tr>
<td>GDD w/ PPE&gt;0.5</td>
<td>0.05693</td>
<td>0.01849</td>
<td>8529.987</td>
<td>9.48</td>
</tr>
<tr>
<td>Heat stress w/ PPE&gt;0.5</td>
<td>-1.68399</td>
<td>0.50826</td>
<td>9881.785</td>
<td>10.98</td>
</tr>
<tr>
<td>High temperature</td>
<td>1.67212</td>
<td>0.65947</td>
<td>5786.9</td>
<td>6.43</td>
</tr>
</tbody>
</table>

Use of GDD/Generations per Year and AVHRR Remote Sensing Models in Combination

In the current study it was assumed that if the host was present in an area the parasite could potentially also be found in the same area. The potential snail distribution, as determined by the AVHRR results, was used as a mask to remove unsuitable areas within the growing degree day model (Fig.12) and the modified heat stress model (Fig.13) for both seasons. The purpose of using the potential snail habitat based on AVHRR data as a mask over the GDD model and the modified heat stress model was to
Fig. 12. Risk assessment maps that combined suitable habitat as determined by AVHRR sensor data and the GDD model. Colored areas represent suitable potential habitat. Darker color tones represent increasing potential severity of *S. haematobium* for (a) wet season (b) dry season.
Fig. 13. Risk assessment maps that combined suitable habitat as determined by AVHRR sensor data and the Modified Risk Index. Colored areas represent suitable potential habitat. Darker color tones represent increasing potential severity of *S. haematobium* for (a) wet season (b) dry season.
show the potential snail habitat and the potential severity of the disease on one map. By dividing results into seasons, the two distributions can be studied to aid in the development or modification of control programs. The distribution, as determined based on AVHRR data, represents the areas that are susceptible during a given time period. The combination of AVHRR determined distribution zone, the mask of unsuitable areas, and the modified heat stress model yielded the most accurate model for *S. haematobium*.

**GDD Model for Water Development Schemes**

When the water balance was not considered to be a limiting factor, the number of potential generations per year ranged from -10 to +18 (Fig.14). The area surrounding Lake Turkana and along the eastern border had the highest number of potential generations starting with 13 and increasing to 18. These regions do not have enough rainfall to sustain a snail population unless water was available through irrigation or other water development projects. The highlands had less than 2 generations which does not support intramolluscan development in the snail host. The area surrounding the highlands along with the Lake Victoria Basin had 5 to 9 potential generations of *S. haematobium* while the coast had 9 to 13 potential generations per year.

**Modified Risk Index Model for Water Development Schemes**

The modified risk index, when corrected for the effects of heat stress on the development of the schistosome, had a risk index range of 0-84 with 0-28 (very low), 29-45 (low), 46-54 (moderate), 55-67 (high), and 68-84 (very high) categories (Fig.15). The areas of very high risk surround the Lake Victoria Basin and the highlands. Moderate and low risk was found around Lake Turkana.
Fig. 14. GDD model that did not use the PPE > 0.5 threshold. By not using the moisture threshold, it is possible to estimate what could happen, if water is made available by man-made environmental change at a given location.
Fig. 15. The modified heat stress model, without the PPE>0.5 threshold, shows the predicted potential severity of *S. haematobium* after introduction of water development projects. The addition of permanent water allows a snail population and *S. haematobium* to develop in new areas.
Discussion

The GDD/Generations per Year Model

The GDD/Generations per year model was based on the effects of temperature conditional on presence of a soil moisture threshold PPE>0.5 on the development of *S. haematobium* within the *Bulinus* snail host. Temperature and moisture are fundamental determinants of the distribution and abundance of a species (Andrewartha and Burch, 1954). In the presence of adequate moisture, thermal regime is the driving force of development, a factor that can be easily measured. Agriculturalists evaluating the suitability of an area for specific crops routinely utilize the GDD method of measuring the rate of development based on temperature above a species-specific ‘base temperature’ below which development does not occur. This method was adapted by Pflüger (1981) to measure the rate of development of *S.mansonii* and *S. haematobium* within their respective intermediate hosts. He described a minimum temperature threshold of 15.3°C for *S. haematobium* in *B.globosus* (base temperature), a maximum temperature threshold of 35°C (thermal death) and a linear relationship of increasing temperature and development rate between the minimum and optimum 22°C temperatures (Fig.16).

Results of the GDD/Generations per year model indicate that areas with two potential generations or less in the snail host have a very low or no chance of schistosome infection in humans. This may be because the amount of time needed to complete this portion of the life cycle takes more time than the life span of the snail (Anderson and May, 1979) due to unfavorable temperature or lack of rainfall or because of temporal relationships of non-consecutive ‘suitable’ periods of the year.
Fig.16. Linear relationship of temperature to the rate of *S. mansoni* development under conditions ranging from the minimum temperature required for development to the optimum temperature of 22°C (Pflüger, 1981).

**Modified Heat Stress Model**

The effect of heat stress on diminishing life cycle suitability at higher temperature regimes, a factor not considered in the initial GGD/Generations per year model, was accounted for in a modified model in two ways: 1) reduced development at temperatures over the optimum for *S haematobium* and 2) the sterilizing effect on *Bulinus globosus* snail hosts at temperatures above 28.5°C.

Pflüger (1981) reported a suitable temperature range for *S haematobium* of 15.3°C to 33°C, with a optimum temperature of development of 22°C and described a hyperbolic relationship between temperature and the length of the prepatent period in the snail, ie. increasing or diminishing development suitability are observed below or above the optimum temperature range (Fig.17). To account for this relationship, a stratified growth curve was developed for the current studies to assign a numerical index value for population. This is supported by field survey prevalence data as compared to 1992/1995 AVHRR annual composite Tmax values extracted from 5 km buffers centered on *S. haematobium* prevalence survey sites by plotting prevalence values from the WHO.
Fig. 17. *S. haematobium* prevalence compared to annual average Tmax from AVHRR annual composite (1992/1995) maps for East Africa atlas (Doumenge et al., 1987) against annual average Tmax over the entire East Africa region (Fig. 17).

The temperature ‘mask’ was developed within the GIS to show both areas and times when monthly mean temperature was >28.5°C for at least 2 weeks during the development period of snail host reproductive capacity. It was assumed that in areas covered by the mask, the snails will either be aestivating or rendered reproductively incompetent by temperature/moisture regime, with resultant ‘null’ transmission potential.

AVHRR Remote Sensing Model

Tmax and NDVI are two factors that can be measured by using AVHRR satellite sensor data to provide surrogate climate parameters that represent thermal-moisture regimes important to the life cycle of snail host-parasite systems. Temperature alone can be a limiting factor for potential snail habitat and no snails were found in areas with temperatures >28.5°C based on the Tmax queries for either season (Fig. 7). However this is not the case for most of Kenya. The majority of Kenya fell within a suitable temperature range for both the wet season and the dry season as determined by AVHRR
query analysis within ArcView 3.3; Tmax queries in fact resulted in an over prediction of suitability in zones considered free of schistosomiasis (Doumenge et al., 1987) (Fig. 8).

The use of NDVI as a surrogate of moisture regime, appeared to be a more accurate limiting factor for snail distribution in Kenya based on GIS query analysis. Suitable NDVI values were found in the highlands and along the coast during the dry season (130-157) and throughout the south and the highlands during the wet season (125-160). There was a difference between the pattern of the wet season and the dry season queries, a reflection of the lack of rainfall during the dry season and the abundance of rainfall during the wet season associated with both the position of the intertropical convergence zone (Wu, 2003) and seasonal patterns of temperature. NDVI alone was not a sufficiently accurate predictor and, like Tmax, over-predicted the endemic area.

In areas selected as suitable for the Bulinus-S. haematobium system using a combination of Tmax and NDVI satellite sensor data, a query based on prediction of the area where the suitability range of both NDVI and Tmax were met resulted in a modeled distribution zone that included 90% of known suitable Bulinus habitat points for the wet season and dry season (Fig. 10). However, the resulting model based on AVHRR derived Tmax and NDVI parameters did not differentiate relative risk within the query-predicted area.

Ecological Niche Modeling

Until recently, most models dealing with the locations of snails and other organisms had assumed a spatially uniform distribution within a selected study area (Anderson and May, 1979). Ecological niche modeling is a method used for finding suitable host habitat that does not assume that the host has a uniform distribution.
Desktop GARP analyzes environmental factors in GIS grid format to create a preferred habitat model. These factors may or may not have an impact on the geographical distribution of a species (Peterson, 2002). Desktop GARP was used in the current study to find suitable habitats for *Bulinus* spp. based on the $5\text{km}^2$ climate grid, and a $5\text{km}^2$ topographic grid taken from the ACT3.0 (Corbett 2002). This created an output of two distributions which describe the wet season and a dry season distribution for Kenya (Fig. 11).

Use of AVHRR `climate surrogate’ data is another means of find suitable host habitat. This method only uses a combination of $T_{\text{max}}$ and NDVI for model development. A statistical program such as SAS or S-Plus must be used to provide statistical validity to this method. Remotely sensed data is easier to collect than climate data or other data used in the GARP program.

**Water Management and Snail Control**

Regional environmental factors used in this study accounted for less than one third (29% or 27%) of the variation of disease risk using multiple-regression analysis (Table 1, Table 2). Humans are able to adapt to a wide range of environments by building shelters or adding water to arid areas. When water is introduced to arid areas, the environments change and water is no longer a limiting factor. The major difference between the GDD model and the modified heat stress model is best illustrated by looking at the two models created for water developments. The GDD model still shows the Lake Victoria Basin as being high risk. The region around Lake Turkana, and the eastern border have become the areas with the highest risk, having 18 potential generations per year (Fig.14). The modified heat stress model maintains that the Lake Victoria Basin is
still the highest risk; the very high risk class has expanded as compared to the model that considers PPE>0.5 (Fig.6) but the overall pattern is the same (Fig.15).

The utility of this modified heat stress model for water development is hypothetical and must wait for validation studies based on field experience. Water management can reduce transmission in areas of flooding by directing the excess water to more arid regions which causes a rapid spread of new habitat and snails (Kloos, 1985). Use of dams to create a constant water level in channels has a great impact on the spread of the snail host and increase of prevalence (Southgate, 2000, Sturrock, 1965). Man-made habitats often provide ideal conditions for the introduction of parasitic diseases in otherwise unfavorable regions (Hunter, 1982). However, snails are able to move vertically in water bodies seeking optimum temperatures (Marti, 1985). This movement allows snails to survive at extreme temperatures in same locations depending on thermal flux dynamics in irrigation water bodies.

Originally a highland model and a lowland model were going to be developed in these studies based on the distribution of *Bulinus globosus* (highland) and *Bulinus physopsis* (lowland) separated by Loker et al (1992). After performing the AVHRR queries, it was discovered that there was no difference for Tmax and very little difference for NDVI requirements for these two species. There was also little difference in elevation between these two species. These two snails are now referred to as a single group because of the similarities between these two *Bulinus* species. Classification of *Bulinus* and other snail species is often controversial due to subspecies and varying strains of the same snail (Kristensen, personal communication). Consideration of both species combined led to the risk assessment maps produced by this project.
Limitations of This Study

A limitation of using 5km\(^2\) climate grid data is that there can be a tremendous amount of variation of both climate and topography possible at regional scales. In highland regions, for example, it is possible to have a difference in elevation of several hundred meters within a 5km\(^2\) area and this difference in elevation affects temperature. For each increase of 620 meters in elevation, the temperature decreases 2.1°C (Hardwick, et al., 1996). Because of this, it is possible in some instances for the GDD/Generations per year model to predict that it is not likely or impossible for *S. haematobium* to develop in some grid cells even though the World Health Organization (WHO) has recorded that the disease was there in earlier surveys.

The use of remotely sensed data may be of limited use for disease prediction when dealing with water development projects. The AVHRR queries for Tmax showed that most of Kenya is suitable for snail habitat and queries for NDVI showed a very limited region that is suitable. The 5km\(^2\) climate grid and the AVHRR sensor are best used on a regional or national scale. To accurately identify high risk zones found within coarse (5km\(^2\)) climate grids, higher resolution sensors such as Landsat TM or SPOT must be used to create a city or village scale risk maps in future work.
Conclusions and Significance

The current work has resulted in the construction of two GIS risk assessment models for *S. haematobium* in Kenya that we propose will be useful in predicting potential risk of disease for control program management in Kenya:

1. **GDD-Heat Stress Index**: This risk assessment method utilizes a GDD-heat stress model, conditional on soil moisture (PPE>0.5), to predict relative risk within the potential seasonal distribution zones defined by AVHRR satellite sensor data (NDVI and Tmax). The model was based on *Bulinus* spp. snail survey records and known development requirements for *S. haematobium* within the current endemic area in Kenya.

2. **GDD-Heat Stress Index for Water Developments**: The GDD-Heat Stress Index was modified to predict hypothetical risk of establishment of *S. haematobium* in future water development schemes, assuming natural rainfall is no longer limiting. The model is the same as the *S. haematobium* suitability model without consideration of the soil moisture (PPE>0.5) limit or AVHRR potential distribution zones.

The GIS models incorporate innovative use of GDD-water budget concepts that are used in crop production models, and apply a unique heat stress modifier methodology to produce digital maps of the suitability of the environment for the *Bulinus-S. haematobium* system in Kenya. Using the power of new GIS/RS tools, data from excellent classical epidemiologic studies can now be systematically placed in an environmental context and used to spatially define the ecological requirements of vector-parasite systems and associated disease risk. Results can potentially be used to develop future in real-time spatial decision support models for health care program management.
References


Vita

Kelsey McNally was born in 1980 in Memphis, Tennessee, to a loving mother Molly McNally and sister Lacey McNally. He was active in the Boy Scouts and achieved the rank of Eagle Scout in December of 1995. In 1998 he graduated from Houston High School in Germantown, Tennessee, and began college that summer. He attended the University of Memphis where he played tuba in the “Mighty Sound of the South” marching band for 2 years. During the summer of 2000, Kelsey taught English as a second language in the Czech Republic and in China that fall. He returned to school in 2001. In the fall of 2001, Kelsey graduated cum laude from the University of Memphis in the fall of 2001 with a Bachelor of Arts degree. While working on a degree in geography, he participated in a study on the erosion effect of channalized streams compared to nonchannalized streams. He then enrolled at the Louisiana State University School of Veterinary Medicine in the spring of 2002 to pursue a Master of Science degree focusing on environmental parasitology. Since his enrollment he as worked on several projects dealing with different diseases in East Africa and Latin America.