UNIVERSITY OF LIVERPOOL

Ecology and Control of Tsetse at the Interface of Conservation and Farming Areas in Northern Tanzania

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Rachel Sarah Lea

October 2019

Acknowledgements

To Prof Steve Torr, I'm not the first student to write that they feel like you took a chance on them, and I'm sure I won't be the last. Thank you for your guidance over the years, both professional and personal, and for the times spent working and learning from you in the field (including how to deal with 'ridiculous' airport situations) in Zimbabwe and Tanzania. Thank you for your unwavering support when I doubted myself, and for always finishing our meetings with a football re-cap. It has been an absolute pleasure and I hope I have repaid your faith.

Thanks also to Dr Alistair Darby and Dr Emily Adams for your supervision and input at the start of my PhD. To Dr Jen Lord, thank you for all your help and hours spent looking over my mostly confused R code, half of this thesis would not have been possible without you. I'm still tired from our Tanzanian HIIT sessions and sorry about the introduction of the daily chips to our fieldwork lunch…

An especially grateful thank you to Jess Lingley, it feels like a long time since you welcomed me to the lab five years ago with your enthusiastic text messages and since then your help has been invaluable. I feel sorry for any future PhD students who will have to get by without you, but I'm sure you'll make a fantastic lab manager, and very well deserved!

To Rhiannon Logan, thank you for your assistance on the insecticide work and for your patience while I remembered how to do basic maths. Prof David Hall and Dudley Farman, those samples took some work! Thank you for helping to explain a complicated process into the language of a non-chemist and for triple-checking everything.

For the crash course in population genetics, many thanks to Dr David Weetman, and for the many hours spent looking over the initially confusing data. Thanks also to Dr Lucas Cunningham for helping select primers and for getting me set up in the dark hood. Thanks also again to Dave, and to Prof Matthew Baylis for your advice and input on my advisory panel.

I have many thanks for all the team in Tanzania at the Vector and Vector-Borne Disease Research Institute and friends and colleagues in Mugumu; Dr Ahmed Lugelo, Benedict Kimbisa, Cuthbert Daffa, Emmanuel Sindoya, Dr Geofrey Mbata, Godfrey Mashenga, Dr Imna Malele, Jackson Mwenda, Judy Lugelo, Dr Mechtilda Byamungu, Dr Oliver Manangwa, Victor Kakenge and Zilpah. I have learnt and grown immeasurably from my time in Tanzania, and you made me feel like I had a new family a long way from home (and thank

you for tolerating my strange cow shaving requests). A special thank you to Dr Furaha Mramba, your enthusiasm and full support for our project most definitely helped it run much more smoothly.

To Dr Harriet Auty, Dr Liam Morrison and Fiona Allan, the times spent working together out in Tanzania are some of those which I look back most fondly on. May there always be one for the road.

Without the support of the ZELS funders none of this would have been possible, I think this programme has been an excellent model for how to support interdisciplinary cohorts of students, and it has introduced me to people and subjects I would never normally have interacted with. Special thanks to Mary Ryan for keeping the cogs turning, and for getting us all to Cambridge, Glasgow, Tanzania and Vietnam for some fantastic meetings. I am also grateful to the funders for the additional money they supplied to allow us to organise the ECR One-Health Symposium. It was an exhausting experience, but one I'm proud to have been involved with, and it was a pleasure doing it with the wonderful Bhagya Chengat, Jennika Virhia and Laura Craighead. Through the ZELS-AS programme I was able to complete an internship in the House of Lords, and to Lord Sandy Trees and Anthony Ridge, I would like to give my thanks for a truly inspiring experience.

Finally, thank you to my family and friends who have been with me on this journey and kept me going with your love, support, and encouragement.

Abstract

Ecology and Control of Tsetse at the Interface of Conservation and Farming Areas in Northern Tanzania

Rachel S. Lea

Rhodesian human African trypanosomiasis (r-HAT) and Animal African trypanosomiasis (AAT) are diseases caused by *Trypanosoma* spp. transmitted by tsetse (*Glossina*). In East and Southern Africa, the risk of r-HAT and AAT is greater in the vicinity of conservation areas where wild reservoir hosts and tsetse are abundant.

The over-arching aim of this study was to analyse fine-scale variations in the abundance and distribution of tsetse at the interface of conservation and farming areas in northern Tanzania. Towards this aim, research was undertaken at the edges of the Serengeti and Saadani national parks to address four specific questions. First, what is the impact of natural environmental variables on the fine-scale (<1 km) distribution and abundance of tsetse? Second, is the treatment of cattle with insecticides by livestock keepers having an impact on tsetse populations? Third, can satellite imagery be used to predict the fine-scale distribution and abundance of tsetse across different agro-ecological zones? Finally, does the molecular genetics of tsetse populations reflect the prior impact of natural and anthropogenic factors on population dynamics.

Analyses of catches of *Glossina pallidipes* and Landsat imagery showed that the abundance of tsetse in conservation areas was correlated positively with normalised difference vegetation index (NDVI), tree cover, soil moisture (Band 7), land surface temperature (LST) and proximity to rivers. In nearby farming areas, the correlations were less marked or not significant, indicating that other factors were controlling tsetse populations.

Questionnaire surveys of livestock keepers showed that in Serengeti district, ~70% of cattle owners reported treating their animals with pyrethroids at monthly intervals. Chemical analyses of hair from cattle using gas-chromatography with mass spectrometry, high-performance liquid chromatography and an insecticide quantification kit showed that ~30% of herds contained animals which had been recently treated with cypermethrin or alphacypermethrin. Previous literature suggests that treatment of cattle with pyrethroids at these levels will reduce the density of tsetse.

Models based on satellite data, developed with data from Serengeti National Park, were successfully used to predict the relative abundance of tsetse in Saadani but were less useful in farming areas where abundance of tsetse was much less than predicted. The low numbers observed may be related to widespread treatment of cattle with insecticide in the farming areas of Tanga region surrounding Saadani.

Analyses of the genetics of *G. pallidipes* populations in Serengeti and Saadani found no evidence for the impact of tsetse control operations. At Saadani, two cryptic and sympatric species of *G. pallidipes* were detected. The discovery of cryptic species may explain intra-specific variation in the behavioural responses of *G. pallidipes* and also makes the use of sterile insect technique a more difficult prospect.

The results are discussed in relation to prospects for improved strategies and approaches to the control of r-HAT and AAT in Tanzania and elsewhere in East and Southern Africa.

Acronyms

AAT – Animal African trypanosomiasis

ANOVA – Analysis of Variance

ASTER – Advanced Spaceborne Thermal Emission and Reflection Radiometer

GCMS – Gas Chromatography – Mass Spectrometry

GDEM – Global Digital Elevation Model

GIS – Geographical Information System

HAT – Human African trypanosomiasis

HPLC – High Performance – Liquid Chromatography

ITC – Insecticide Treated Cattle

IQK – Insecticide Quantification Kit

LST – Land Surface Temperature

LSTM – Liverpool School of Tropical Medicine

MMR – Mark-Release-Recapture

NDVI – Normalised Difference Vegetation Index

NRI – Natural Resources Institute

PCR – Polymerase Chain Reaction

r-HAT – rhodesian human African trypanosomiasis

VVBDRI – Vector and Vector-Borne Disease Research Institute

WHO – World Health Organisation

For figures: TL – Top left, TR- Top right, LL – Lower left, LR – Lower right

Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 African trypanosomiasis

1.1.2 Rhodesian human African trypanosomiasis incidence and geographical distribution

The 27 reported cases of r-HAT in 2017 are recorded from seven countries which provided data to the World Health Organisation (WHO). In order of case numbers, these countries were Uganda (13), Malawi (7), United Republic of Tanzania (3), Zambia (3) and Zimbabwe (1) (Figure 1). Kenya and Rwanda had no reported cases.

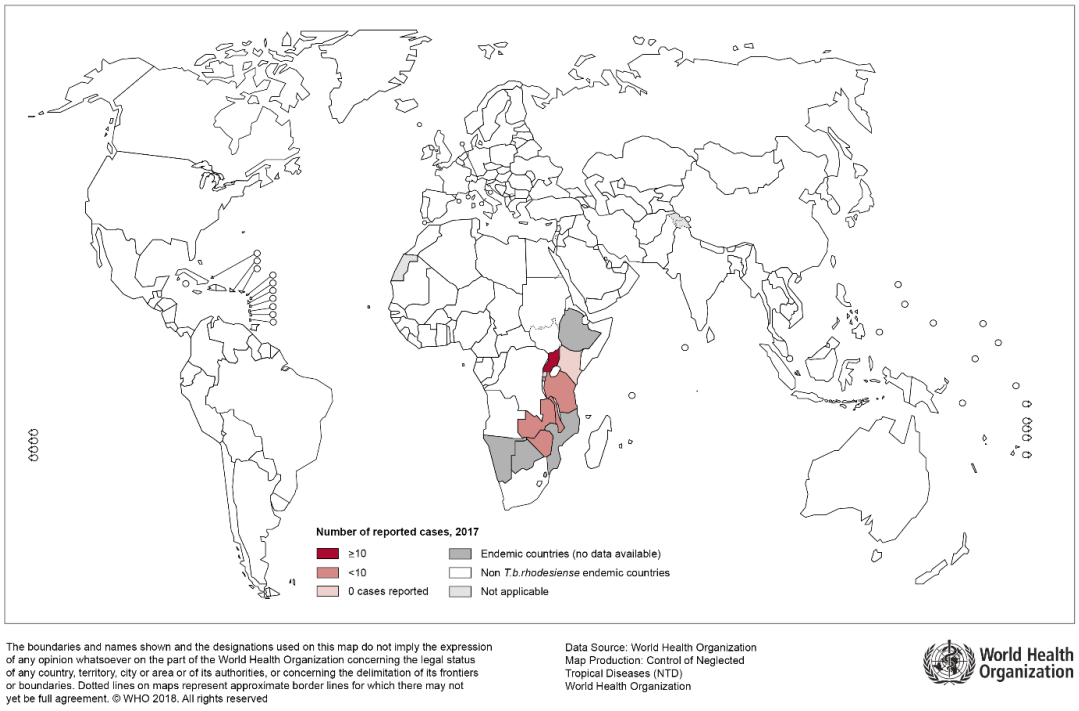

Figure 1. Distribution of r-HAT in 2017 (unadapted (WHO, 2018)*).*

Reported case numbers are however likely affected by under-reporting, as the disease generally affects the poor in remote rural areas (Kuepfer *et al.*, 2011) and is often misdiagnosed as HIV, malaria or TB (Lejon, Jacobs and Simarro, 2013). In Uganda it is estimated that for every one r-HAT fatality, there are a further 13 deaths which go unreported (Odiit *et al.*, 2005) and there is a predicted 12.3 million people in Africa at risk of becoming infected (Simarro *et al.*, 2012). The number of cases has dropped considerably over the last 25 years due to control efforts. For example, tsetse and trypanosomiasis were eliminated in the Okavango Delta of Botswana with aerial spraying (Kgori, Modo and Torr,

2006) and from large areas of Sebungwe in Zimbabwe through a combination of host elimination, ground spraying and aerial spraying (Hargrove, 2003). The rate of decrease has however slowed and has been at a low level for several years because for many areas the costs of a vector control operation at an appropriate scale are too high.

Rhodesian HAT foci are closely associated with wilderness areas (Figure 2) where there are large numbers of wild hosts, vectors and appropriate vegetation which allows tsetse to persist. Farming is restricted in these areas, so there is little habitat degradation, maintaining the environmental conditions that allow tsetse populations to thrive. Foci associated with wilderness areas are also the areas where tourists have been infected whilst visiting the National Parks and Game reserves. Over 1 million tourists visited Tanzania in 2012, and provided direct employment for over 400,000 people, contributing to 9.9% of GDP in 2013, with the Serengeti National Park being a major draw (World Bank Group, 2015).

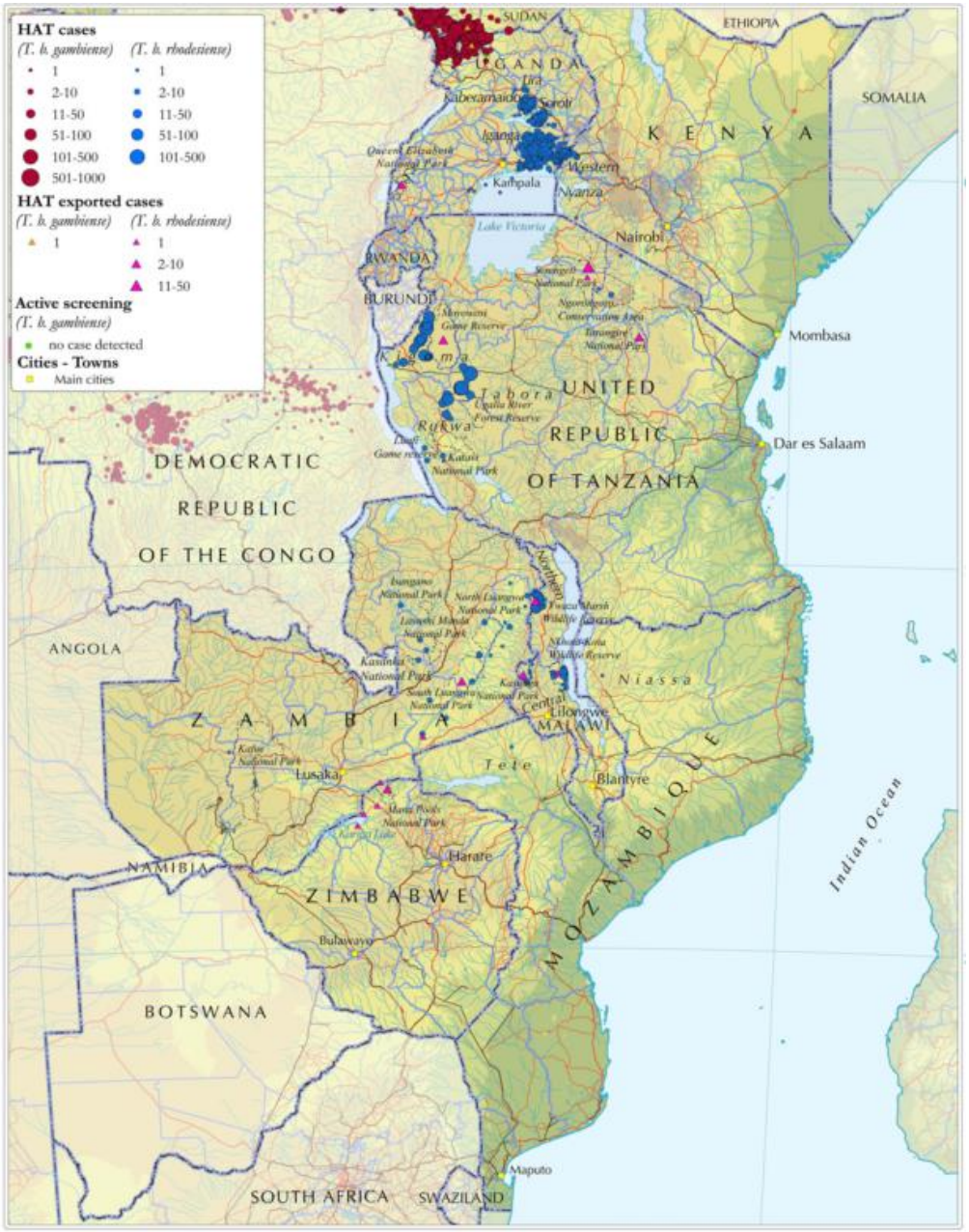

Figure 2. The location of human African trypanosomiasis cases which were reported between 2000-2009 in Eastern and Southern-eastern Africa. Rhodesian HAT cases are in blue, and the number of cases identified by circle size in the key. Exported cases are those which have been diagnosed elsewhere but attributed to the highlighted area (adapted from (Simarro *et al.*, 2010)*).*

1.1.3 *T. b. rhodesiense* life cycle

Transmission of r-HAT is related to both the life cycle of the *T. b. rhodesiense* parasite and the tsetse fly. When a tsetse takes a blood meal from an infected host to gain the energy and nutrients necessary to survive, they ingest bloodstream trypomastigote forms which travel with the blood meal into the fly midgut (Figure 3). The trypanosomes then

differentiate into procyclic trypomastigotes, begin to replicate and exit the midgut through the peritrophic membrane into the proventriculus. Here they differentiate into first mesocyclic trypomastigotes and then into epimastigotes before travelling to the salivary gland through the oesophagus, proboscis and hypopharynx. In the salivary gland they continue to multiply and develop into metacyclic forms. It takes between 18 and 35 days for the trypanosome to reach this stage in the tsetse and thereafter the fly will remain infectious until it dies. If the tsetse feeds on a host once the trypanosomes have established themselves in the salivary glands, they will be injected into the host's capillaries and lymphatic system, where they will develop into long slender forms before being transported to the blood stream via draining lymph nodes. Once in the blood stream they can be found as either a long 'slender' form which can multiply, or as a 'stumpy' bloodstream trypomastigote which will be ready to transform into the procyclic form in the fly if it is taken up in a blood meal. The bloodstream forms enter a variety of bodily fluids such as the lymph and cerebrospinal fluid.

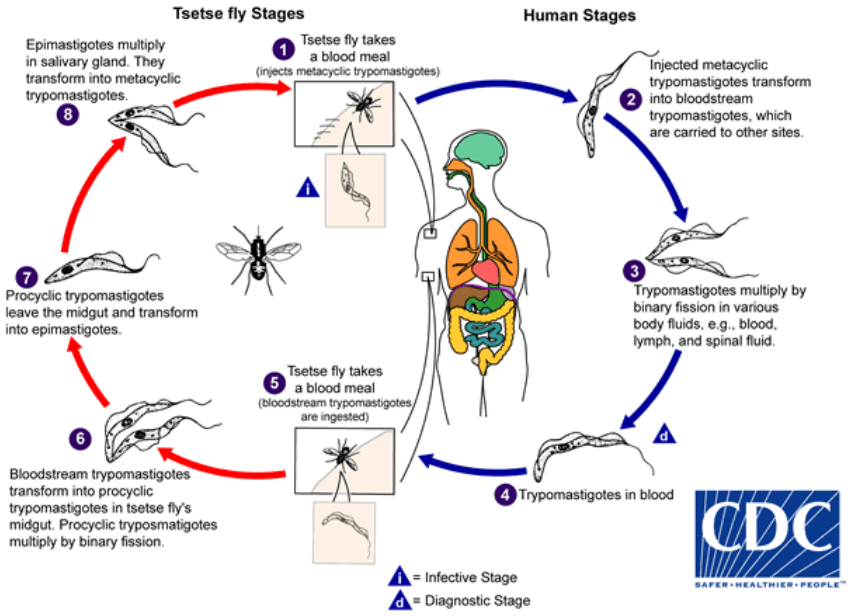

Figure 3. Trypanosoma brucei life cycle through a tsetse and human host (unadapted from CDC (2015)).

1.1.4 Symptoms

The infection in humans has two stages and each has a different set of symptoms. Following introduction by the bite of the fly the parasites circulate in the blood stream and lymphatic system of the host. During this first stage the symptoms include a chancre on the skin around the bite, headaches, fever, weakness and joint pain. When the parasites cross the blood-brain barrier and invade the central nervous system the second-stage symptoms

begin. The victim will suffer from worsening confusion and a variety of neurological conditions before entering a coma which will, in the majority of cases, lead to death in the absence of intervention. Rhodesian HAT is an acute disease, with the progression to death from the onset of symptoms taking around six months (CDC, 2012).

1.1.1 Human African trypanosomiasis

Human African trypanosomiasis (HAT), commonly known as Sleeping Sickness, is a parasitic disease caused by trypanosomes which are transmitted by the bite of a tsetse fly (*Glossina* spp.). HAT is a neglected tropical disease which is responsible for millions of human deaths since the early 1900s. People most at risk of the disease are those living in rural areas who rely on animal husbandry, fishing and agriculture, often in places with poor infrastructure and healthcare access. There are two forms of the disease; Gambian HAT (g-HAT) occurs in West and Central Africa and is caused by the parasite *Trypanosoma brucei gambiense*, and Rhodesian HAT (r-HAT) occurs in East and Southern Africa and is caused by *T. b. rhodesiense*.

Both forms of the disease are generally considered to be fatal if left untreated. However, r-HAT is an acute, faster progressing disease leading to death within months of infection whereas for g-HAT, death occurs 2-3 years after infection. Humans are the main reservoir of g-HAT, whereas the reservoirs for r-HAT are predominantly wild animals and, in some areas, cattle.

Due to extensive control efforts, the numbers of g-HAT cases have decreased from approximately 26,000 in 2010, to 1420 new reported cases in 2017 (WHO, a). Whilst this is a large decrease and probably the lowest global incidence for >100 years, there is a risk that global incidence could increase following, say, civil unrest or socio-economic crises. The majority of the new cases in 2017 were in the Democratic Republic of Congo, which is particularly unstable (Tong *et al.*, 2011). Reported cases of r-HAT have also decreased since 2000, from approximately 700/year to 27 new reported cases in 2017 (WHO, b). The majority of r-HAT cases were found in Uganda, which is also currently the only country where both forms of HAT are endemic.

1.1.5 Diagnosis

Diagnosing the stage of disease is important as there are different treatments for each. During the first stage of the disease, infection can be detected via microscopy of a blood sample. Once the infection has progressed to the second stage the trypanosomes can no

longer be detected in the blood and a lumbar puncture must be performed to examine the cerebrospinal fluid for presence of trypanosomes.

Active screening of the human population is commonly used for the control of g-HAT, the form of sleeping sickness found in West and Central Africa, as a method of disease management (Welburn, Molyneux and Maudlin, 2015), however the progression of r-HAT is too rapid for mass screening to be an effective means of control.

1.1.6 Treatment

Due to the variant surface glycoproteins which cover the surface of the trypanosome and can randomly change during each generation, the prospects for developing a vaccine are small (Vickerman, 1985). Chemotherapy is therefore the only current treatment and the drugs for r-HAT were developed many decades ago and have significant, life-threatening side effects themselves.

1.1.6.1 Suramin

Suramin is used to treat first stage r-HAT. It was first introduced in 1922 (Nok, 2003) and it is thought to work by inhibiting binding of glycolytic enzymes in trypanosomes (Barrett *et al.*, 2007) and is given as a single injection once a week for six weeks (Gastellu Etchegorry *et al.*, 2001). It cannot however cross the blood-brain barrier and is therefore not used in the second stage of the disease (Bacchi, 2009). Serious side effects include adrenal failure and exfoliative dermatitis (Gastellu Etchegorry *et al.*, 2001).

1.1.6.2 Melarsoprol

Melarsoprol is used for treating second stage r-HAT. This is an arsenic-based drug which is able to cross the blood-brain barrier. It is also thought to interfere with glycolytic enzymes (Bacchi, 2009). It is given as three injections every 7-10 days for three cycles. Serious side effects include myocardial damage, exfoliative dermatitis and encephalopathy. Encephalopathy is reported in 5-10% of patients and when it does occur, there is only a 50% chance of survival (Gastellu Etchegorry *et al.*, 2001).

1.1.7 Animal African trypanosomiasis

Trypanosomes not only cause a human disease in Africa, they are also responsible for animal African trypanosomiasis (AAT) in livestock. AAT is caused by several species of *Trypanosoma* including *T. brucei brucei, T. vivax* and *T. congolense.* These species are not pathogenic to humans but can cause death and/or production losses in most species of livestock, estimated to cause annual losses of billions of US dollars across Africa (Meyer *et*

al., 2018). In cattle for instance, AAT reduces meat and milk production, capacity for ploughing, and increases mortality and abortion rates (Swallow, 2000).

The distribution of the disease has left many parts of Africa unsuitable for farming for both livestock production but also for crop production, as the breeds of cattle more appropriate for use in ploughing are those with little trypanosome tolerance. Cross breeding between trypanotolerant and susceptible species and has been undertaken to attempt to offer a solution to this problem (Orenge *et al.*, 2012).

Unlike the human disease, animals can be protected against infection with the use of prophylactic drugs for AAT. The only current prophylactic drug is isometamidium chloride, which can also be used as a therapeutic. This, along with diminazene aceturate, a drug which cures the disease, are the two most common treatments used in Africa, however there are six licenced compounds. Resistance of trypanosomes to trypanocides is increasing and is a concern for future sustainable management of AAT (Giordani *et al.*, 2016).

1.2 Tsetse ecology

1.2.1 Tsetse life cycle

As underreporting of r-HAT is widespread and the treatments can have very serious side effects, the most cost-effective means of preventing the disease is to target the vector and hence reduce the transmission in the first instance. Especially in the case of r-HAT which is a zoonosis, treating the disease in humans has little impact on transmission. For effective vector control to take place, the life cycle of the tsetse fly must therefore be understood.

Amongst insects, tsetse have an unusual reproductive cycle, known as adenotrophic viviparity. Females are mated when they are about five days old and about a week after mating, she produces one egg which will develop into a larva whilst still in her uterus. Nine days later the L3 larva is deposited onto the ground and buries itself into the soil. Once in the ground it pupates and emerges as an adult roughly 30 days later (Leak, 1998). As mean life expectancy is 1-2 months, a female is likely to produce four offspring in her lifetime (Franco *et al.*, 2014).

1.2.2 Tsetse groups

There are three types of tsetse, grouped according to their natural habitat. The first is the riverine group ('Palpalis-group') which is found in West Africa and is associated with rivers and lakes systems. The second is the forest group (Fusca-group), which is found in a range

of forest types from dry forest islands to riparian vegetation in savannah areas to rainforests. The third; and group of interest for this study, is the savannah group (Morsitans-group) of tsetse, and its distribution is closely linked with savannah woodlands and the wild animals associated with these ecosystems.

1.2.3 Savannah tsetse ecology

Savannah tsetse distribution is correlated closely with the distribution of wild animal hosts and, as their name suggests, presence of savannah woodland. However, within wilderness areas different species of the savannah group have slightly different habitat types and seasons where and when they are relatively abundant. In West Africa, *G. m. submorsitans* abundance peaks early in the dry season, and puparia are often found in dry soil within isolated forest areas or riparian vegetation. *G. longipalpis* has a smaller distribution than *G. m. submoristans* and is generally associated with riparian vegetation. In East Africa, *G. pallidipes* also occupies riverine habitats within savannah woodlands but is more tolerant of hotter and drier conditions due to its relatively large size and thicker walled pupae. *G. austeni* is the smallest of the savannah tsetse and is restricted to the East African coast and associated with areas that have high rainfall and dense vegetation (Leak, 1998). *Glossina swynnertoni* has one of the smaller distributions of the savannah group, found only in northern Tanzania and Kenya. It is most often caught in more open areas associated with *Acacia* woodland and pupae are found in open thickets (Ngonyoka, Gwakisa, Estes, Nnko, *et al.*, 2017).

1.3 Tsetse control

As the reproductive rate of tsetse is very low, only 4% of female flies need to be killed a day to eradicate a population (Hargrove, 1988). However, as the larvae burrow into soil to develop they cannot be targeted using the larval source management used against mosquitoes. Consequently, all control methods are directed against adult tsetse.

1.3.1 Methods of control

1.3.1.1 Historical tsetse control

Methods of controlling tsetse have changed and developed over the last 100 years when tsetse were first incriminated as the vectors of sleeping sickness. Before the use of insecticides, common methods were host elimination and bush clearing. Host elimination was initially very effective as the source of blood meal for the flies was removed, however these areas would then be used by farmers, often with livestock which provided re-invading

tsetse with a host. Bush clearing was also very successful as this removed any suitable habitat for the tsetse, however was very environmentally damaging (Malele, 2011).

1.3.1.2 Ground spraying

Once insecticides such as DDT and dieldrin were discovered in the 1950s, ground spraying was introduced. The insecticides were sprayed on sites which tsetse rest, such as the bases of trees. The development of persistent organochlorines meant that sprayed surfaces were lethal to tsetse for several months, long enough to target flies after they had emerged from pupae. From 1955 up to 1964, a ground spraying operation with DDT was undertaken in Nigeria to eradicate tsetse in the Chad river system. The project was very successful and cleared 230,000 km^2 of tsetse. While ground spraying was important, the very dry seasons were believed to have helped as this reduced suitable tsetse resting places so spraying could be applied selectively (Davies, 1964).

1.3.1.3 Aerial spraying

Aerial spraying was developed in the 1970s and involved the use of insecticides such as endosulfan and deltamethrin, which were released from planes as a mist to target very large areas (Leak, 1998). One of the first reported uses was in Zambia using endosulfan for five applications with three week gaps between each over 1600 km^2. Eradication of tsetse from their target area was achieved, however reinvasion occurred due to the absence of an effective barrier to prevent reinvasion of tsetse from neighbouring, unsprayed areas (Park *et al.*, 1972). It was also initially successful in the Lambwe Valley of Kenya supporting ground applications after the 1980 outbreak of r-HAT using endosulfan and pyrethrum. Nine applications of endosulfan reduced *Glossina pallidipes* trap catches by >99% and following this, three applications of pyrethrum were made however this further application did not impact tsetse populations. The effort was however not coordinated well with ground operations and tsetse populations rebounded, and large quantities of endosulfan fell into the watershed of Lake Victoria and toxicity to fish populations had not been considered (Wellde *et al.*, 1989). More recently, aerial spraying was used to eliminate tsetse and trypanosomiasis from the Okavango Delta Botswana (Kgori, Modo and Torr, 2006).

1.3.1.4 Sterile insect technique

The sterile insect technique (SIT) involves releasing irradiated flies to compete to mate with the healthy population which introduces sterility and causes the population to gradually decrease. The single example of this technique being used to eliminate tsetse and trypanosomiasis is on Zanzibar, where tsetse were completely eradicated from the island.

Over 8 million gamma-sterilised male flies were released on to the island over 3 years from 1994, and due to its island setting providing a natural barrier to reinvasion, tsetse have been absent from the area ever since (Vreysen *et al.*, 2000). While SIT can be very effective under certain conditions, it is a very expensive method and requires large barriers against reinvasion (Torr, Hargrove and Vale, 2005).

1.3.1.5 Artificial baits

The development of artificial baits came through an understanding of how tsetse locate their hosts. It was discovered that tsetse used both visual and olfactory stimuli therefore it was recognised that simulating natural hosts with traps or targets (large insecticide-treated panels of cloth) baited with artificial host odours could be used to control tsetse (Vale, 1974). Savannah tsetse species were found to be attracted to large blue and black horizontal objects with odours (Torr *et al.*, 2011). The most effective odours are acetone, 1-octen-3-ol, 4-methylphenol and 3-n-propylphenol (Torr *et al.*, 1997).

Targets have been used extensively since their development in the 1980s. In 1984, basic black cloth and netting targets where used in a control effort in the Zambezi Valley of Zimbabwe covering 600 km^2. At the end of a six-month period, tsetse populations at the centre of the control area had reduced by 99.99%. An area of 300 km^2 was covered in Zambia between 1989 and 1991 also using the black cloth and netted target and reduced tsetse densities by 94.8% after 3 months. Blue targets were also used to support the SIT effort on Zanzibar (Vreysen *et al.*, 2000). In South Africa, blue and black targets were used between 2001 and 2004 to reduce *G. austeni* populations in the Greater St. Lucia Wetlands Park. *G. austeni* females densities were reduced by 99% after 13 months (Esterhuizen *et al.*, 2006). The downside of targets is that over time the colours become faded due to dust and sun exposure and therefore become less attractive. The insecticide also degrades over time and becomes less effective at killing tsetse (Vale *et al.*, 1988).

1.3.1.6 Insecticide treated cattle

Another method of control is the use of insecticides on livestock. As flies attempt to take a blood meal from cattle, they will come into contact with the insecticide on the surface of the animals and will then die before they can feed on another host, so will no longer be able to transmit trypanosomes. This is often a very cost-effective option (Kajunguri *et al.*, 2014) in areas where there are many cattle, and is also useful in reducing tick-borne disease as many products used to control tsetse are also effective against ticks.

1.3.1.7 Recent control efforts in other African countries

In Botswana, the method of aerial spraying was used to target tsetse over a 16,000 km^2 area of the Okavango Delta (Kgori, Modo and Torr, 2006). Deltamethrin was sprayed on five occasions for a total of eight weeks during 2001 – 2002 and a barrier of deltamethrin treated targets placed between two treatment blocks. Pre-treatment catches from man fly-rounds were 44.6 – 100, while surveys conducted for three years after treatment did not catch any flies and it was stated that elimination had been achieved. Whilst eliminating tsetse from the area has reduced the need and cost of tsetse control in the future, the initial costs for aerial spraying are very high, and likely prohibitive for many African countries.

Uganda is another country where aerial spraying was employed (Welburn *et al.*, 2006) however this was not able to eliminate tsetse in infested areas after treatments in 1988 and 1990. In this time there was little information exchange between livestock keepers and policy makers about effective control strategies, but policy makers were not getting the epidemiological information they would have needed to give appropriate advice (Brightwell *et al.*, 2001).

In 2006 a campaign called Stamp Out Sleeping Sickness was initiated as a public-private collaboration, with the aim of treating 86% of cattle with trypanocides in five districts in a buffer zone between *T. b. rhodesiense* and *T. b. gambiense* infected areas (Welburn *et al.*, 2006; Bardosh, Waiswa and Welburn, 2013). This successfully reduced trypanosome prevalence in cattle by 75%. This progress was, however, damaged by movement of untreated cattle into the districts. The programme was then modified from a top-down approach of providing mass treatment, to a bottom-up approach by placing five veterinary graduates into communities with the aim of increasing knowledge, providing local services for treatment and setting up spray teams. It was established that pyrethroid-based insecticides would also be needed to help deal with tsetse re-invasion and the entry of un-treated cattle from other areas. A study by Bardosh, Waiswa and Welburn, (2013) examined the use of insecticides in these districts using questionnaires and interviews between 2011 and 2012. Use of insecticide was reportedly high, with 70.5% of farmers using insecticide monthly during the rainy season, however only just over half of people were using insecticides effective against both ticks and tsetse. Even in districts which were high-risk for HAT, the volume of product effective against ticks only was double that of products effective against both. The reasons for this included a poor understanding of the disease and vectors, brand recognition, price, mode of action, product residual period,

availability and poor information dissemination. Even when farmers knew that the amitraz products available in the shop were more expensive per animal if made up to the recommended dilution, the actual product in the shop was often cheaper to buy and this made it more appealing. Shop attendants did not frequently provide information on the most appropriate product to use, and so farmers would often pick the cheapest one or a product they recognised. It was suggested that market regulation would be the best way to increase the number of cattle treated with pyrethroids effective against tsetse (Bardosh, Waiswa and Welburn, 2013).

1.4 Modelling of tsetse control

The abundance and age structure of vector populations are important determinants of vectorial capacity and this is reflected in the Ross-Macdonald model of malaria transmission (Macdonald, 1956) amongst other epidemiological models of vector-borne diseases. The original Ross-Macdonald model was adapted by Rogers, (1988) for sleeping sickness and this approach has been adopted by others over the last 30 years (Rock *et al.*, 2015). These models have been used to inform general approaches to controlling the disease. The Ross-Macdonald model for instance highlights the importance of vector control which not only reduces the abundance of adult mosquitoes but also the mean age of the population. Reducing these two parameters has a much greater impact than say larval control. Rogers' model of African trypanosomiasis highlighted the small contribution of human hosts to the transmission of human African trypanosomiasis which implied that control efforts directed against the animal hosts and/or tsetse would be more effective. More recently, models of sleeping sickness have been used to assess the impact of various control strategies on Gambian sleeping sickness (Rock *et al.*, 2015; Mahamat *et al.*, 2017). None of these models consider spatial variation in the distribution, abundance and age structure of tsetse populations.

Similarly, early models of the population dynamics of tsetse did not consider spatial variation in abundance (Langley and Weidhaas, 1986; Hargrove, 1988) or considered movement separately from growth (Hargrove, 1981; Rogers, 1997). The first efforts to model the movement and growth of a tsetse population in a two-dimensional framework utilized the Fisher equation applied to a grid of cells to simulate the impact of various control strategies (Hargrove, 2000; Hargrove, Torr and Kindness, 2003). These models were implemented using Visual Basic for Applications implemented within the Following on from this approach, Vale and Torr, (2005) produced a model ('Tsetse Muse') which explicitly modelled the movement and growth of tsetse along a one-dimensional transect. The

mortality of tsetse within cells along this transect could be varied, which allowed simulation of spatial variation in population mortality, but not growth or movement. This approach was used to model tsetse control operations using aerial spraying (Kgori, Modo and Torr, 2006), insecticide-treated cattle (Torr and Vale, 2011) and targets (Vale *et al.*, 2015). Developing this approach still further, Vale developed a two-dimensional cellular automata programme which modelled variation in movement and growth of the tsetse population and also included variation in host composition and vegetation across a 50 x 50 km landscape. The programme also attempted to model the transmission of trypanosomes between humans, livestock and wildlife. The programme was the first to attempt to model the effect of different vegetation types (e.g., riverine woodland, thicket, grassland) on the movement, growth and mortality of tsetse. The model ('HAT-trick') is available at www.tsetse.org but no publications arose from this work.

An alternative approach to modelling the dynamics of tsetse and trypanosomes across a landscape has been the use of agent-based models (Muller, Grébaut and Gouteux, 2004). These mechanistic models aim to simulate individual tsetse ('agents') as they move, feed, reproduce and die across a landscape where vegetation and hosts can vary. With both the cellular automata and agent-based models, there are a large number of assumptions about the effects of, say, vegetation on the movement and growth of tsetse. While the inputs to the agent-based models are sometimes derived from specific locations (Alderton *et al.*, 2018), the aim of the models is to derive broad principles rather than simulate the dynamics of tsetse and trypanosomes at a specific site.

There has been a long-standing recognition of an association between vegetation and the dynamics of tsetse populations (Ford, Nash and Welch, 1970; Hargrove and Vale, 1980). Using spatial information to guide tsetse control operations was established in the 1960s when ground spraying was widely used. Planners used aerial photographs to identify woodland areas where spray teams should apply DDT to the trees where tsetse rested during the day. The development and widespread availability of satellite imagery, geographical information systems and powerful computers has allowed the development of geostatistical models to analyse relations between vegetation and tsetse.

Geostatistical models can be used to examine the relationship between tsetse abundance and the environmental covariates related to each tsetse trap. Within the framework of the model it is acknowledged that in the trap count data there will be a spatial dependence which cannot be accounted for by the selected environmental variables. These models have

been used recently to examine tsetse abundance in Tanzania (Lord *et al.*, 2018) and Uganda (Stanton *et al.*, 2018).

1.5 Remote sensing

In the past 25 years, many studies have examined the relationship between habitat and the distribution and abundance of tsetse using remotely sensed data from satellites. Values for covariates in geostatistical models (1.4) are often gained from remote sensing. Remote sensing describes how cameras and sensors on satellites detect different wavelengths of reflected and emitted radiation from Earth to define the physical attributes of an area. Combining values gained from satellite imagery and catch data from traps can assist with identifying the ecological factors which contribute to differences found in abundances within areas.

There are several different types of satellite imagery which can be used for this purpose. The Landsat program has been gathering images for over 40 years, with the most recent satellite, Landsat 8, in operation since 2013. The resolution of Landsat 8 images is 30 m and it measures eight spectral bands including red, green, blue, near infrared and thermal infrared (Barsi *et al.*, 2014) with a revisit frequency (the time taken for the satellite to return to same point to take a repeat image) of 16 days. From the Sentinel program – the Sentinel-2 satellites have been in operation since 2015 and offer 13 spectral bands, a revisit time of five days and a resolution down to 10 m (Li and Roy, 2017). Planet Labs have been launching small satellites since 2015 and now have over 175 orbiting the earth (GIS Geography, 2019). It measures four bands; red, green, blue and near infrared at a 3 m resolution and has a daily revisit frequency. The Terra satellite launched by NASA in 1999 has a sensor called the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) which calculates a Global Digital Elevation Model (GDEM) at approximately 30 m resolution (Hirt, Filmer and Featherstone, 2010). Tsetse have been shown to have defined geographical limits based on elevation (Leak, 1998). For other environmental variables, Landsat imagery has been used in many tsetse research projects due to the availability of images over long time-scales.

Landsat images are located using a ‘path’ and ‘row’ system; with the paths as almost vertical tracks running from east to west, and rows running from north to south. Within the Landsat 8 spectrum, there are several bands which can be used to examine factors which are likely to influence suitability of tsetse habitat. These factors include soil moisture (Band 7: 2.11 -2.29 µm) (Kitron *et al.*, 1996), normalised difference vegetation index (NDVI)

(Bands 4: 0.64 – 0.67 µm & Band 5: 0.85 – 0.88 µm) (Ngonyoka, Gwakisa, Estes, Salekwa, *et al.*, 2017) and land surface temperature (Band 10: 10.6 – 11.19 µm) (Lord *et al.*, 2018). Soil moisture is measured in images from the amount of reflectance of water from the soil; it is likely that areas with a higher soil moisture are cooler and more shaded, providing vital resting places to protect from desiccation (Jackson, 1954). NDVI is essentially a measure of the greenness in an area. Healthier, more green vegetation will absorb more visible light and reflect less than less healthy, browner or more sparse vegetation. The more vegetation in an area, the more likely it is that there will be more shade for resting adults, soil at a temperature and humidity that suits developing pupae and higher concentrations of wild (e.g., buffalo, warthog) and domestic (e.g., cattle) hosts.

1.6 Tsetse movement and distribution

Infected cattle can be responsible for spreading *T. b. rhodesiense* into previously HAT-free areas (Fèvre *et al.*, 2005), however the movement of tsetse is also involved with disease spread over short distances (Aksoy *et al.*, 2013). The presence of hosts will determine if and how tsetse can move across large areas, however habitat suitability is also important. Dispersal will be greatest when the flies move into areas where they can gain a fitness advantage (Schreiber and Lloyd-Smith, 2009), which is likely to be areas with lower human density due to a decreased degradation of vegetation (Wardrop *et al.*, 2013). There is a balance between the benefit of flies being in open areas to find hosts visually and therefore not having to rely solely on olfactory senses (Kagbadouno *et al.*, 2011), and being close enough to thick vegetation that they can return to for shelter from the hottest parts of the day to avoid desiccation (Brightwell, Dransfield and Williams, 1992).

A knowledge of the movement and distribution of tsetse in an area is vital for planning operations to monitor or control tsetse. There are several ways to measure the movement of tsetse, using both non-molecular and molecular methods.

1.6.1 Mark-release-recapture

A theoretically simple non-molecular method of measuring movement is mark-release-recapture (MRR). In this process, flies are caught, marked with an identifying feature such as paint and released. Traps set across the area will then be monitored to determine if any marked flies have been recaptured and their original trap location compared with where they were re-caught to determine how far they have travelled (Hargrove, 1981). There are difficulties with this method, as due to the distances that flies travel each day few flies are ever recaptured, and hence it is only suitable in areas with high densities of flies. There are

also ethical issues regarding the release of flies which could spread disease (Solano *et al.*, 2009).

1.6.2 Genetic markers

More recently it has been more common to employ molecular methods for measuring movement, through the use of markers such as mitochondrial or microsatellite DNA. While costly due to the laboratory work involved, this removes the need for multiple trips to collect and re-collect flies, is less reliant on high densities of flies, and there are no ethical concerns as the flies will be killed.

1.6.2.1 Method differences

MRR and population genetics studies often show very large differences in their estimates of fly movement, with ecological methods generally showing higher rates of movement than genetic approaches. For instance, Ouma, Marquez and Krafsur (2005) suggested that flies could not be travelling anywhere near 1 km a day (Vale, Hursey, Hargrove, *et al.*, 1984) as genetic differentiation was much higher than expected between populations which could have theoretically reached each other within their lifetimes. Vreysen *et al.*, (2013) however states that population genetics can only give general indications of movement, and that results should be backed up with MRR.

There is also a discrepancy between estimates of the rates of movement of male and female tsetse. Koné *et al.*, (2011), Ouma *et al.*, (2011) and Mélachio *et al.*, (2015) all report that males move more than females whereas MRR-based studies show the opposite (Vale, Hursey, Hargove, *et al.*, 1984). Mélachio *et al.*, (2015) suggests that females may be travelling greater distances day-to-day, but may be returning to the same site to oviposit, however there is no other evidence to suggest this is true.

1.7 Thesis objectives

The Serengeti National Park is an important historical focus of sleeping sickness. Its establishment as a conservation area was in part because of the health problems posed by high densities of tsetse. The parks in the Serengeti ecosystem (Serengeti, Tarangire, Maasai Mara) area are still the major source of exported HAT cases in East Africa, and there is likely to be many unreported local cases. Although numbers of HAT in Tanzania are generally low (<10 cases/year), this is the time when maintaining control is most important. As cases reduce and less attention is brought to a disease, less funding may become available for control. Control must therefore be as cost-effective as possible and must be sustainable for communities to be able to maintain it.

Findings from Lord *et al.*, (2018) show that in the Serengeti study area numbers of tsetse were declining rapidly outside the borders of protected areas. This study builds on this finding by examining how the density and distribution of tsetse changes over a finer scale. In particular, this study aims to quantify the environmental and anthropogenic factors underlying this decline. Extending the work of Lord *et al*., (2018), this study also examined how environmental factors and human activity affected tsetse population in coastal areas of Tanga region with its different farming systems and climate.

To achieve these aims, the studies present in this thesis aimed to address four specific questions using three different approaches, to provide complementary evidence. The four questions are:

1. Do environmental variables alone explain variation in the distribution and abundance of tsetse over fine scales?

2. Are the livestock-management practices of livestock keepers in Tanzania also affecting the distribution and abundance of tsetse and thereby controlling tsetse through their own actions?

3. Can predictive models of tsetse abundance developed in one region to address the first question, be successfully applied to other parts of Tanzania?

4. Can population genetics methods be used to quantify the impact of ongoing control operations or assist in their design?

Chapter 2: Impact of environmental factors on the fine-scale distribution and abundance of *Glossina* species on the edge of the Serengeti National Park

2.1 Introduction

Tsetse can travel up to 1 km a day (Vale, Hursey, Hargove, *et al.*, 1984) and are often modelled moving diffusely (Bursell, 1970) which largely ignores the effect of habitat (Hargrove, 2000). Studies have however shown that numbers of flies within study areas do not appear to disperse equally (Odulaja *et al.*, 2001), especially in the case of *G. pallidipes*. Odulaja *et al.,* (2001) found that although tsetse species did appear to be randomly distributed during a period of short rains, during both the dry and long rainy seasons, flies were aggregated in certain areas. The level of aggregation was attributed to maximum temperature. Brightwell *et al.,* (1992) found high numbers of *G. pallidipes* associated with the thick vegetation found along the edges of rivers in Kenya, with similar patterns observed in the Serengeti National Park (Lord *et al.*, 2018). It is not necessarily however the presence of the river which is directly contributing to an increase in the number of tsetse. It could be more likely due to the increased amount of vegetation present, the type of vegetation present, the soil moisture levels or the temperature.

If it could be determined which of these factors, singly or in combination, contributed to an increased abundance of tsetse, geostatistical models could be created at a fine resolution (~100 m) to accurately predict areas of high risk. Knowing where the highest numbers of tsetse will be found within an area would also be useful for several aspects of control, such as gaining a clearer indication of the realistic population size to help with monitoring a control operation. Applying control using traps or targets will also be most effective when targeting areas with the highest densities of flies.

There is one study by Brightwell *et al.*, (1992), which examined the distribution of *G. pallidipes* and *G. longipennis* along a transect from rivers into open areas. Spread of both species from the rivers was recorded to be at least 3.5 km. *G. pallidipes* were found most commonly no further than 500 m from the riparian vegetation surrounding the river and it was during the rains when flies were caught up to 3.5 km away. Age structure of flies was compared between the three vegetation categories found along the transect - riverine thicket, open acacia woodland and open plains. Generally, there was little difference in the age distributions, however for category 0 flies (i.e., about 0-8 days old), there was a significant difference in that the highest proportion of these flies were found in the riverine

thickets, and the lowest proportion in the open plains. Uterine contents, which indicate the stage of larval development, were examined and most categories were found to be similar. Significant differences were only found with flies containing a second instar larvae as the percentage of these flies in the open plains was significantly lower than in the other vegetation types. No significant difference was found between the distribution of males and females from the river despite evidence that females disperse more than males (Vale *et al.*, 1985).

There are many papers more generally examining the relationship between savannah tsetse, vegetation type and other environmental variables. One of the first studies examining the effect of habitat on tsetse flies is from Jackson (1945) who introduced species of tsetse into areas more commonly associated with different species. For example, *Glossina morsitans* and *G. palpalis* were introduced into the habitat of *G. swynnertoni* and *G. pallidipes* and it was found that *G. morsitans* were capable of producing offspring however *G. palpalis* were not. Laboratory experiments showed that *G. swynnertoni* were less resistant to weight loss of both fat and water at 25.3°C than *G. morsitans*, suggesting that *G. swynnertoni* may need more humid habitats. The overall results highlighted different physiological requirements between species in their habitats.

Environmental conditions are also considered to impact on the distribution of tsetse flies. Rogers (1979) used fly round records of *G. morsitans submorsitans* from Nigera to calculate density independent mortalities and correlated them with average annual temperatures and saturation deficits (the amount by which the water vapour in the air must be increased to achieve saturation without a change in temperature or pressure). The author was able to establish an environmental optimum of 24-26 °C and 5-13 mm Hg saturation deficit. This paper and Rogers & Randolph (1986) concluded that tsetse species were sensitive to the temperature, rainfall and the saturation deficit of an area.

2.1.1 Satellites, GIS and models

In the past 25 years, many studies have examined the relationship between habitat and the distribution and abundance of tsetse using remotely sensed data from satellites. Rogers & Randolph (1991) began using satellite images (at 25m^2 resolution) to compare NDVI with the environmental factors they had already been studying. They found that NDVI was highly correlated with saturation deficit, positively (but not linearly) correlated with rainfall, and not with temperature. NDVI was then plotted against mean catches of *G. palpalis* and *G. tachinoides* and highly negatively correlated results were found, demonstrating the ability of remote sensing to predict abundances of flies.

Satellite images were again used by Kitron *et al.,* (1996) to investigate the distribution of *G. pallidipes* in Kenya. In this case, a finer resolution (30 x 30 m) from Landsat 5 was used and soil moisture measured by Band 7 was found to be mostly highly correlated with fly densities. Stepwise multiple regression was also performed and it was found that when Band 7 was combined with Band 3 (linked with chlorophyll absorption) and Band 6 (thermal band used in vegetation stress analysis and soil moisture discrimination) the R^2 value for densities recorded in 1988 was 0.88, showing that 88% of the variation in count data could be attributed to these factors.

Robinson *et al.*, (1997) used the factors maximum temperature, average temperature, minimum temperature, rainfall and NDVI to attempt to characterise differences between the species *G. morsitans centralis*, *G. morsitans morsitans* and *G. pallidipes*. The NDVI resolution was large at 7.6 km however as this analysis was across several countries in southern Africa the scale is not unreasonable for this purpose. It was found that for *G. m. centralis* the most accurate predictor was the average NDVI with 75% correct predictions, for *G. m. morsitans* it was the maximum of the minimum temperature with 84% correct predictions, and for *G. pallidipes* it was also the maximum of the minimum temperature with 86% correct predictions. Robinson *et al.*, (1997a) took this analysis further and gained even more accurate predictions after sub-dividing the habitat into two groups. The first were habitats which were cold and the elevation was low, and the second was where it was hot and the elevation high.

Risk of sleeping sickness to villages in Uganda was calculated by Odiit *et al.*, (2005) using satellite-derived variables along with the village population numbers, the distance of the village to a hospital capable of diagnosing sleeping sickness. The satellite variables included NDVI and soil moisture and a buffer area of 1.5 km was used. No actual fly densities were used however the presence or absence of sleeping sickness cases was included. The buffer used was large however for a country-scale to be able to establish which villages were more at risk than others this would be sufficient. It would not however be able to assist with the identification precisely where tsetse control should be implemented. Symeonakis *et al.*, (2007) performed analysis to identify areas which should be the highest priority for tsetse control across Zambia. They identifed areas under pressure from agriculture and presence of cattle and identified areas with high absorption capacity which was defined as low cropland use intensity and stocking rates, and not likely to suffer environmental degradation under low soil erosion rates. High priority areas were designated are areas

with a high absorption capacity which were next to areas of high land pressure. This again was on a country-wide scale.

Gondwe *et al.*, (2009) examined the distribution and density of *G. m. morsitans* at the interface of the Nkhotakota Game Reserve in Malawi. Land cover maps were generated using satellite data with a 20 m resolution (accuracy 88%) and the vegetation within each 20 m block classified into three groups. The first group was woodland with trees covering more than 50% of the surface, down to group three which was sparse vegetation. Buffer zones of 200 m around each trap were set and the proportion of each vegetation class within the zone calculated. Tsetse abundances were found to increase with increasing distance from the boundary into the reserve by a factor of 1.35 per km, for each percentage increase in the proportion of vegetation belonging to group one inside each buffer, the tsetse abundance increased by a factor of 1.02.

More recently, the effects of habitat fragmentation on tsetse populations have been the subject to analysis. Ducheyne *et al.*, (2009) set 238 transects to catch *G. m. morsitans* in Eastern Zambia and overlaid hexagons onto satellite images of the study area, at both 250 m and 500 m sizes. Fragmentation was calculated within each hexagon by measuring the amount of 'natural vegetation' – miombo and munga vegetation types and 'disturbed vegetation' – agriculture and village land. Hexagons were then classed into five categories according to the proportion of each type of vegetation within them. They found the majority of flies were caught in non-fragmented areas, and as fragmentation increased, a decline in abundances was observed.

Geostatistical models are increasingly being used to predict the distribution of tsetse. Matawa *et al.*, (2016) used Landsat TM data at 30 m resolution from Zimbabwe to analyse potential habitat suitability for *G. pallidipes*. They found high probabilities (>50%) that tsetse would be found between NDVI results of 0.3 and 0.6 and the suitability of habitat decreased further away from drainage lines.

Catch data from the Serengeti National Park was recently used with Band 7, NDVI, Land Surface Temperature (LST) and elevation in a model to predict decline of tsetse numbers away from the protected area (Lord *et al.*, 2018). Values for the environmental variables were averaged within a 500 m buffer around each trap from traps placed ~ 1.5 km apart. A predictive model was created using data from traps placed > 10 km inside the Serengeti National Park to ensure there was no external factors which may have been influencing catch. Band 7 was found to be the variable most related to *G. pallidipes* abundance,

followed by NDVI, LST and finally elevation. The predictive model generally predicted a decline in the abundance of *G. pallidipes* with distance from the protected area and into farmland. There were however areas where the model overpredicted catches on the border and in farming areas. Entomological surveys in these areas did not catch many tsetse whereas the model predicted tsetse to be abundance. For *G. swynnertoni,* the geostatistical model suggested a more even distribution of tsetse across habitats but still with a decline in farming areas.

2.1.2 Relationship between landcover and abundance of tsetse

Van den Bossche & de Deken (2002) investigated the distribution of *G. m. morsitans* in eastern Zambia in relation to cattle and vegetation type. Transects and MRR experiments were performed in an area with two main vegetation types. The first was miombo, an open woodland with *Brachystegia* and *Julbernardia* trees no higher than one storey, and the second is munga, a one or two storey height woodland with *Acacia, Combretum* and *Terminalia* trees. Catches in miombo peaked during the rainy season and were lowest during the cold and dry seasons. The opposite was observed in munga. The catches of tsetse in both areas were highly correlated with the abundances of cattle from the previous month, and MRR highlighted flies moving between the two areas which accounted for sudden increases in one area which could not be explained by climatic conditions.

Tsetse do not always appear to behave how we expect. *G. brevipalpis* is historically thought to be confined to vegetation with a dense overhead canopy to provide plenty of shade with a high humidity. Esterhuizen *et al.* (2005) however caught large numbers in open grassland and in exotic plantations. The authors were unsure whether these were flies breeding in these other areas, or whether they had travelled out of the dense vegetation to feed. This is epidemiologically important as it had previously been thought this species was only found in areas which were difficult for humans or livestock to access, however they could now be playing a much greater role in the transmission of HAT.

Sciarretta *et al.* (2010) examined the spatial clustering of *G. m. morsitans* and *G. pallidipes* in Ethiopia. Both species were found to be aggregated along a main and smaller rivers, and within savannah areas which had heterogenous vegetation. There were no large catches found in open savannah located far from water.

2.1.3 Nutrition and age of flies

Vegetation may have an impact not only on apparent density of tsetse but also their physiological condition. The effect may be apparent over large distances (e.g. >1 km)

between habitat types. Savannah tsetse require a habitat with a plentiful supply of hosts to feed upon, enough vegetation to provide shade to avoid desiccation from high temperatures, but not so much vegetation that the visual host-seeking ability would be impaired. Different vegetation types can cause conditions such as humidity and temperatures to vary. As mentioned previously, tsetse flies are highly mobile vectors and are predicted to move in a diffuse manner (Hargrove, 2000), so it would seem unlikely that they would be in one place for long enough for the quality of the habitat to impact on the nutritional state of the fly. However, if there was a noticeable impact on the condition of the flies in certain habitats then this would have implications for control as healthier flies are more likely to live longer and travel greater distances – increasing the risk of transmission of disease.

Factors which can be measured to determine the condition of the flies include the fat and haematin content. Haematin content is measured to determine how long it has been since the last blood meal. The nutrients from the first two or three meals are used to build flight muscle (Bursell, 1961), however after this the amino acids are used to begin to produce fat (McCabe and Bursell, 1975). Fat builds up over the course of two days and then begins to decrease (Figure 4) and it can be used to infer the level of success the fly has had at gaining blood meals over the last few hunger cycles (Jackson, 1946). These processes however generally relate to male flies as females use most of their fat for their larvae.

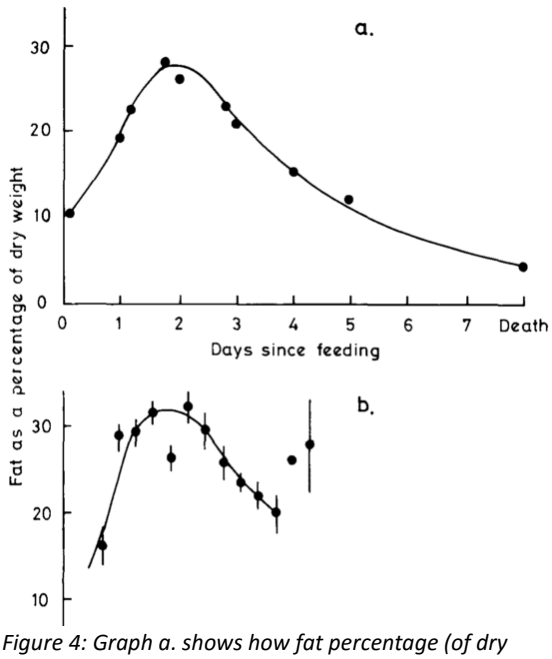

Figure 4: Graph a. shows how fat percentage (of dry weight) varies after a blood meal, and graph b. shows how the log of haematin varies with the fat percentage. Data for both graphs is obtained from male G. m. morsitans (Adapted from (Hargrove & Packer, 1993)).

The difference of conditions of flies in different habitats would be particularly important around the edges of wilderness areas, as the habitat will vary greatly from outside the areas, across the boundary and into the park. This could also be important as some areas along the boundaries may have more suitable vegetation than other areas which would enable the flies present to remain fitter, and be more likely to spread further out of the wilderness areas into human settlement areas and therefore increase the risk of transmission of disease.

Rogers and Boreham (1973) examined fat content of *G. swynnertoni* in the Serengeti National Park. A Land Rover was used as a moving bait and flies were caught with hand nets. Three areas were sampled; area one was mostly open with the exception of several trees and shrubs along dried river beds, area two had an abundance of *Acacia* and *Commiphora* trees and there was also several dried river beds and area three had a much greater quantity of vegetation and had the largest range of animals. Area three was also much more likely to retain surface water after the other two areas had dried out. Fat content for male flies was very similar across the three areas however was higher in area three (4.80 mg compared to 3.31 mg and 3.71 mg in areas one and two respectively), which had the greatest catch size. The pattern was similar for the female flies, although the difference in fat content was larger with area three again having the greatest values. Post-teneral females were excluded from the study as the larvae will contribute to the fat values. The catch size in area three was greater than in area two however was equal to area one. The difference in fat content in these areas was attributed to the greater abundance of hosts in area three, as flies would be able to find meals much more frequently and easily.

Pilson & Pilson (1967) conducted a study to examine the feeding and resting behaviours of *G. m. morsitans* in Zimbabwe to discuss the relevance of discriminative clearing as a control method. It was found that in the nine different habitat types they sampled in with a stationary ox bait, there did not appear to be a concentration of flies in any particular vegetation throughout the year.

Bursell (1966) considered that the implications from the results of the Pilson & Pilson (1967) study warranted further investigation due to the impact it could have on control recommendations (1967 data gathered before 1966 paper was published). The author decided to examine *G. morsitans* in four vegetation types – *Brachystegia* woodland, *Colophospermum* mopane woodland, riverine fringe and *Combretum* bushland. In each vegetation type there were two trapping methods employed. The first method employed was a human fly-round carried out between 3pm and 6:30pm, and the other was with an ox as a stationary bait between 5:30am and 7pm (only flies which probed were collected). Fat content was measured by weighing the flies before and after treatment with chloroform. The length of the hatchet cell was also measured to give a proxy of mean size, the level of wing fray was recorded, and the quantity of residual blood-meal. The size of the flies is important as this is dependent on the nutritional state of the parent female and therefore provides an indication of the condition of the female population in the 1-2 months prior to the catching date (Glasgow & Bursell, 1961). Haematin and fat content were measured. The

results suggested that flies which were caught in *Brachystegia* and *Colophospermum* mopane woodlands had the highest haematin and fat content values, followed by those found in the riverine fringe and then by those found in *Combretum* bushland (Bursell, 1966). These results were surprising to the author as the vegetation in the *Colophospermum* mopane woodland was entirely leafless and appeared to be very unsuitable for tsetse as there was no shade and day time temperatures which reached almost 40°C. The *Brachystegia* woodland had slightly more suitable vegetation however conditions still appeared harsh. In contrast, the vegetation in the riverine fringe was in full leaf and shade was readily available however the nutritional state of flies caught in this area was much poorer than in the *Brachystegia* and *Colophospermum* mopane woodlands. It was only in the *Combretum* bushland that nutritional state appeared to match expectations, as the vegetation was mainly long grass and condition of the flies was poor. Wing fray was greatest in areas where haematin and fat content was low, and it is suggested that this is due to an increased time spent hunting for a host, rather than it indicating that the flies are older which might normally be assumed. The size differences of the flies was significantly different between sites, suggesting that the populations were isolated even though they were only several miles apart.

Fragmentation of habitat was examined with relation to the age of *G. m. morsitans* and proportion of females in Zambia by Mweempwa *et al.* (2015). Only the ages of females were established as the ovarian dissection method was used. In Lusandwa, the least fragmented area, the age structure of the population was skewed so that there were more young flies than old in all seasons. In Chisulo, the most fragmented area, the opposite was observed and more old flies were found than young in all seasons. In Zinaka, a moderately fragmented area, mostly older flies were observed during the dry seasons, and mostly younger flies were found during the rainy season. The authors concluded that in Chisulo, the high fragmentation leaves the area unsuitable for breeding and the old flies present are ones which have travelled longer distances and dispersed into the area. Lusandwa is likely to be suitable for breeding in all seasons and the population in Zinaka during the dry season consists of mainly dispersing flies, but is suitable for breeding during the wet seasons. The proportion of female flies varied depending on fragmentation. The highest proportion of females were found in the least fragmented areas, and the lowest proportion in the most fragmented. This was attributed to the least fragmented areas having more possible breeding sites for females.

As well as ovarian dissection, the pteridine method can also be used to determine the age of tsetse flies. Ovarian dissection requires skill, is time-consuming, and cannot determine age when the fly has already been through four ovarian cycles. Pteridine is an aromatic chemical compound which accumulates linearly in the head of a tsetse as it ages, and can be quantified due to it's fluorescent properties (Lehane and Mail, 1985).

2.1.4 Chapter objectives

As recently established, large numbers of *G. pallidipes* have been caught at or very near rivers in the Serengeti National Park area (Lord *et al.*, 2018). Rivers form a fundamental part of life in rural sub-Saharan African communities as they are used for collecting drinking water, watering livestock, bathing and clothes washing (Franco *et al.*, 2014). It is therefore important that the dynamics of tsetse in these riverine areas are well understood, as these areas could be hotspots for disease transmission due to the increased contact between people and flies, and to estimate how far flies may travel from rivers depending on the surrounding vegetation to assess the risk to people living and working nearby.

Analysis of remotely sensed environmental variables has not yet been undertaken in the Serengeti National Park area with relation to *G. pallidipes* and *G. swynnertoni* abundances at a fine scale (~100 m). While Lord *et al.*, (2018) examined tsetse predictions at a 500 m resolution, their traps were placed relatively far apart (~1.5 km) so would not highlight if there was clustering of populations within a small area.

From the literature it appears that the abundance of similar species do not display similar correlations with variations in landcover. Other factors, such as host density or vector control operations may also affect the abundance of tsetse over relatively short distances.

Many studies have analysed relationships between environmental variables estimated from satellite imagery at relatively coarse (>1 km) scales. This chapter describes studies that aimed to look at relationships between the abundance of tsetse at finer (<1 km) scales. Such a study will contribute to our understanding of where people and livestock might be at greatest risk of infection by tsetse-borne trypanosomes.

There have also been very few studies which have directly studied the effect of vegetation and habitat type on nutritional status, and those which have shown either no difference or have found results which appear contradictory to expectation. Understanding the physiological status of tsetse in different habitats will also contribute to our understanding of why, where and when tsetse bite humans.

Accordingly, the objectives of studies presented in this chapter are to investigate:

1. How the abundance of *G. pallidipes* and *G. swynnertoni* changes with distance from river.
2. The environmental factors which cause changes in abundances.
3. Age structure and nutritional status of tsetse across different habitats

2.2 Materials and Methods

2.2.1 River transects - February 2016 sampling

To measure abundance and movement of flies in association with river systems, diagonal transects were performed at specific locations. The first trap was placed as close to the river as possible, and the following traps were placed ~200 m apart and each ~100 m away from the river (Figure 5). This was repeated on the other side of the river in the opposite direction. The traps were deployed in a diagonal rather than an orthogonal transect to reduce the possibility that traps closer to the river caught flies that would have otherwise been caught in traps further away.

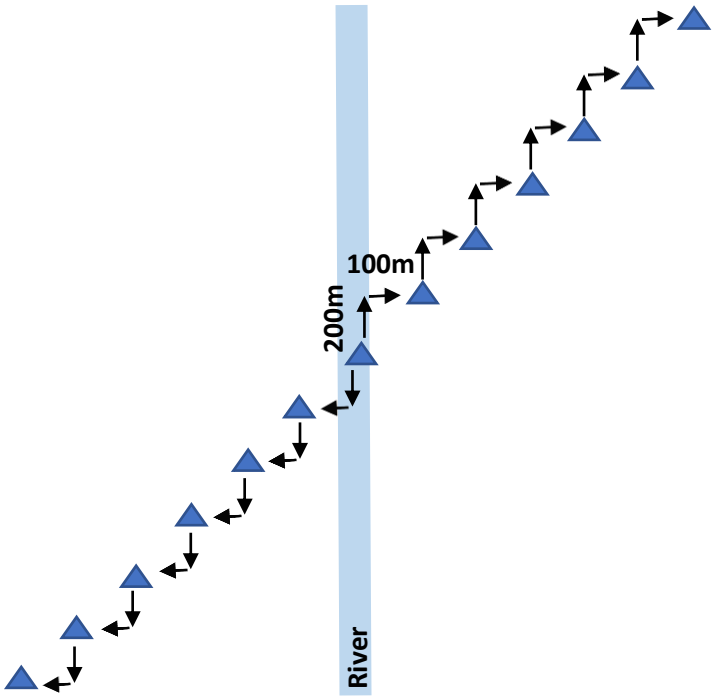

Figure 5: Diagram of how the transects were positioned along the river. Each blue triangle represents a trap.

The four locations were picked specifically to provide contrasts between each other, and in some cases on either side of the river (Figure 6).

1. The Grumeti transect has farming area on one side of the river and game reserve on the other showing a contrast of good and bad habitat on either side.
2. The Park Nyigoti transect was on one side in the Ikona wildlife management area where habitat was expected to be suitable for tsetse, and has grazing habitat on the other which is expected to be less suitable. There is however another game reserve approximately 1 km on the other side of the grazed area.
3. The North Ikorongo transect has both sides of the transect in the game reserve however, one side of the river has slightly degraded habitat as it is near the edge of the reserve.
4. The last was near Robanda, on the south-westerly edge of the Ikorongo game reserve, which borders the Ikona East wildlife management area.

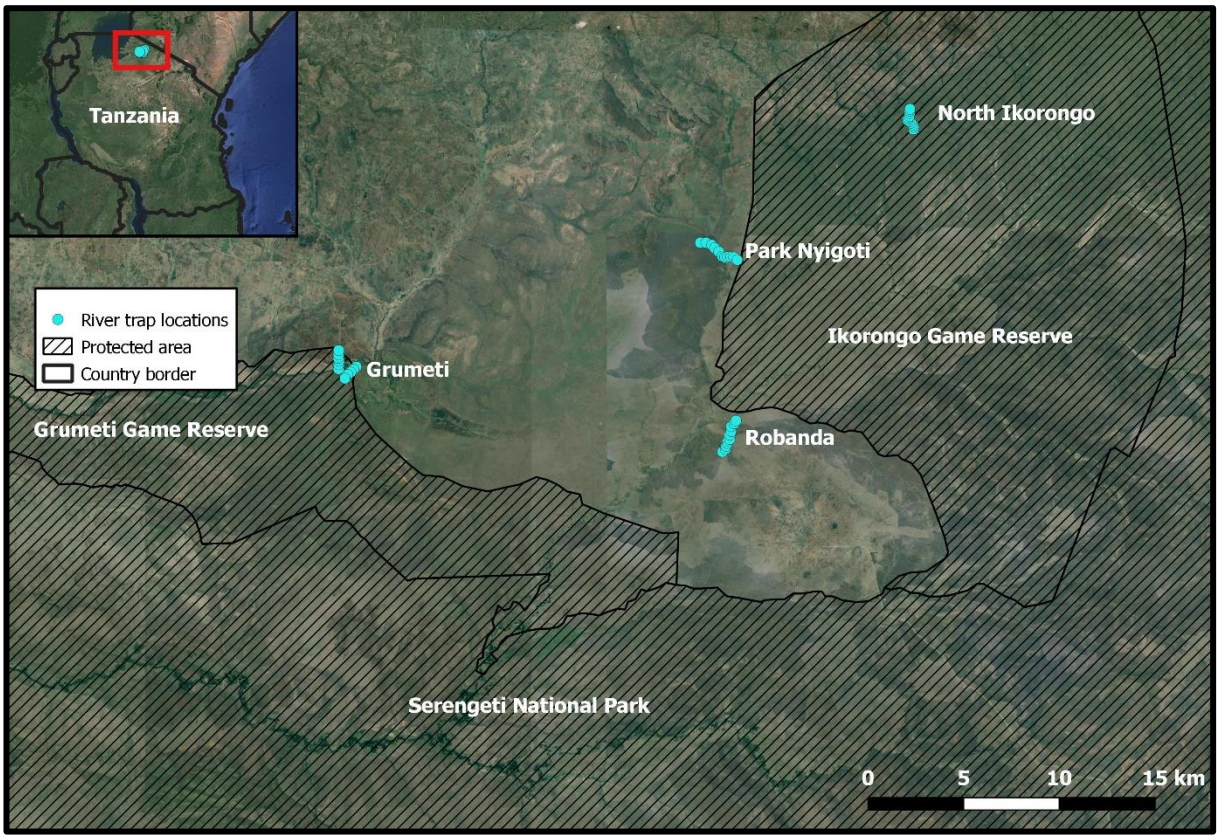

Figure 6: Map showing the location of the four study sites where transects were performed in February 2016. Each blue circle indicates a trap position.

Traps were set for five days and collected at 24 h intervals. Traps were baited with acetone (500 mg/h), 1-octen-3-ol (0.4 mg/h), 4-methylphenol (0.8 mg/h) and 3-n-propyphenol (0.1 mg/h) dispensed using the methods of (Torr *et al.*, 1997). Flies caught in these traps were counted and sorted to measure abundance at each site. The catches were analysed to assess whether abundance was related to environmental variables. Average catch was

calculated by taking the log of the combined male and female catch of each species for each trap, and an average taken of the log values for each day. This averaged log value was mostly used in analysis and back transformed to show on maps by performing log^10 +1. A linear model was fitted to the log-transformed average counts, each time with a single environmental covariate as a predictor.

2.2.2 Grid sampling - October 2016 sampling

Following the transect survey conducted in February 2016, a more extensive and detailed field study was performed in October 2016 to evaluate the effects of environmental variables on abundances of tsetse. Instead of a single line of traps, a grid system was employed. This was so a larger more representative area could be covered. Four sample sites were selected (Figure 7), two of which covered the North Ikorongo and Grumeti transects from January 2016. Between 15 and 18 traps were placed across each area covering approximately 1km^2. Two of these areas – Grumeti and north Ikorongo, were selected due to their location on the border. Inner Ikorongo and Ikoma were selected due to their location being ~ 10 km inside a protected border so would not be influenced by anthropogenic factors. The shape of the grid in each area was determined by logistical and safety factors, such as the density of vegetation and proximity of wild animals. Sites were selected that contained a range of vegetation densities within each grid. Grids were set for three days and traps collected every 24 h and baited as for the February 2016 sampling.

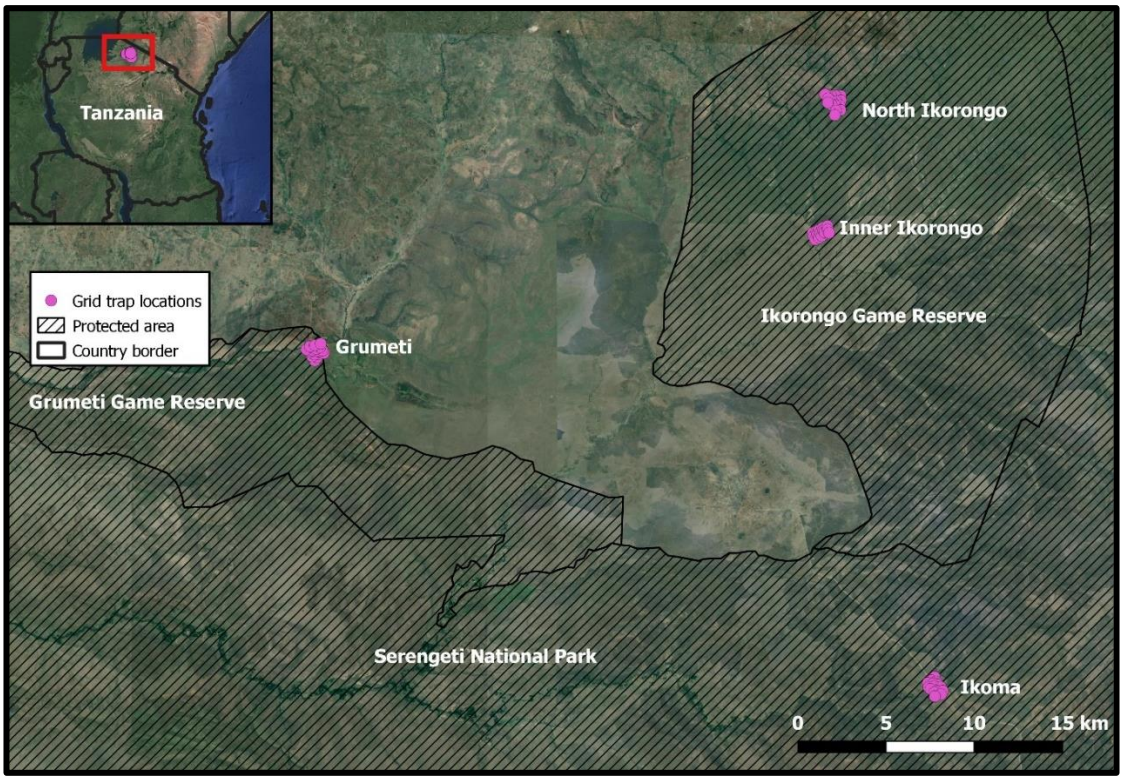

Figure 7: Map showing the location of the four study sites where the grids of traps were placed in October 2016.

2.2.3 Satellite image processing

Environmental variables NDVI, Band 7, LST and elevation were examined using Landsat 8 and ASTER Global Digital Elevation Model (GDEM) images. The study sites were split across satellite path 169 row 62 and path 170 row 61 so two images needed to be taken for each time point. For the river transects conducted in February 2016, the image for path 169 row 62 was taken on the 16/02/2016 and the image for path 170 row 61 was taken on the 23/02/2016. For the grid sampling conducted in October 2016, the image for path 169 row 62 was taken on 20/10/2016 and the image for path 170 row 61 was taken on 29/10/2016.

The Landsat 8 images were processed by first cropping so that any edges which have no data and contain zero values were converted to NA values. Clouds were then removed from images and new images created using the QGIS plugin 'Cloud masking for Landsat products'.

Each image had an associated metadata file, and a file with a subset of the data containing only the data relevant for the required band was created. Landsat 8 images are already converted to radiance so only needed to be converted to reflectance. Landsat 8 images are converted to reflectance using top of atmosphere values and values of the sun elevation (Appendix A.3.3). The first image to process is for Band 7. This requires only a conversion to reflectance. As NDVI requires both Band 4 and Band 5 images, once the images are

converted to reflectance the two images are overlaid and a function is applied with the formula: (band.5 - band.4) / (band.5 + band.4).

To process temperature a Band 10 radiance raster must be combined with an emissivity raster which is created from the NDVI raster (Sobrino, Jiminez-Munoz and Paolini, 2004). Using NDVI, a new raster is created called emissivity which accounts for the reflectivity of different land surfaces. This uses NDVI to assign new values where NDVI is less than 0.2 (0.973) because this is likely to be bare soil, and greater than 0.5 (0.99) which is indicative of a fully vegetated pixel. Another value (0.986) is given for NDVI values between 0.2 and 0.5 which are composed of a mixture of bare soil and vegetation. This is done in the QGIS raster calculator (Appendix A.3.6). The emissivity raster created from Step 2 with the raster calculator is then used with the Band 10 radiance raster to create the final temperature raster. The elevation raster is downloaded from NASA Earth Data with the required data set the ASTER GDEM which has a 30 x 30 m resolution. For Band 7, NDVI and LST, the images from the two different paths were then histogram matched to equalise the images and account for any differences in sun angle or brightness, before the mosaic function (raster package) in R joined them together with a mean value taken for any overlapping pixels.

Distance of each trap from the river was calculated using a downloaded river shapefile (ICPAC GeoPortal) and processed in R against GPX points. Percentage tree cover was gained by downloading the Hansen Tree cover image (Hansen *et al.*, 2013) for both the river transects and grids, and a manual method was also employed for the grids by drawing polygons around trees in QGIS. The images used in QGIS are from Google Earth with a resolution of 3 m and dated from 2014 and 2015. The Hansen tree canopy cover file shows cover from trees taller than 5 m in height and is dated from 2010. Trees were selected manually if the canopy was larger than 10 m in diameter, and the percentage of the total area of the polygons within a 100 m radius buffer zone was calculated. Vegetation smaller than 10 m was likely to be shrubs or trees less suitable to provide shade to tsetse (Figure 8). The images used is QGIS for tree cover and the Hansen image are from several years apart and may not be comparable, however using larger trees should reduce the risk of seasonal changes impacting on results. Large-scale events such as forest fires between 2010 and 2014/2015 may however impact the similarity between the images.

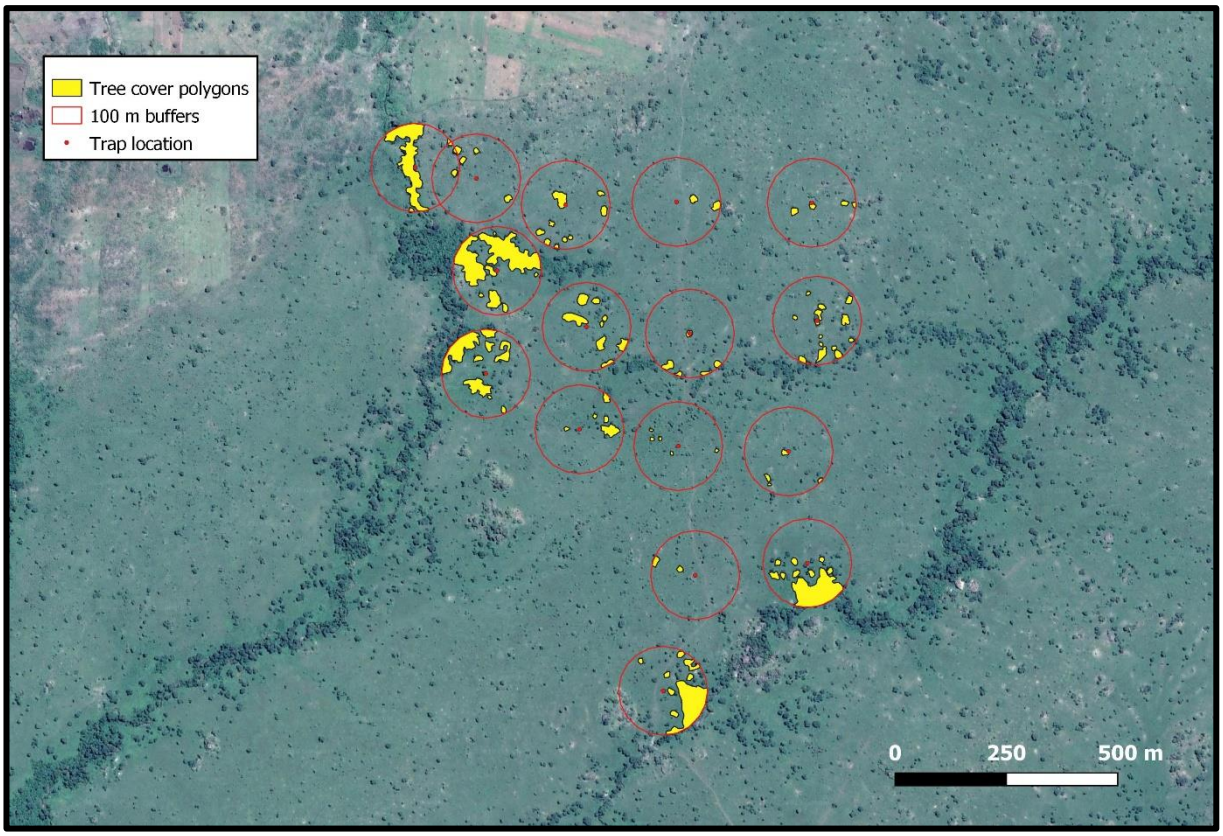

Figure 8: Example from North Ikorongo of how polygons were placed to calculate percentage tree cover.

To obtain average values for Band 7, NDVI, LST, elevation and percentage tree cover (Hansen), a buffer zone of 100 m radius was applied around each trap and the average taken of all the pixels within the buffer. All of the environmental variables were analysed in a series of generalised linear models with log-transformed catch as the dependent variable and the environmental variables as potential explanatory variables. Average catch was calculated as in 2.2.1. Significance was assessed through the P value (<0.05) and the adjusted R^2 value which describes the proportion of the variation in the dependant variable (e.g. abundance) which is explained by the independent variable (e.g. NDVI).

2.2.4 Estimating the age of tsetse

Flies caught in each grid were counted and sorted to measure abundance. *G. pallidipes* female flies (n=63) and *G. swynnertoni* female flies (n=64) were examined by ovarian dissection, and the heads of 1091 flies were stored for pteridine analysis to determine their age (Lehane and Mail, 1985). Only flies which were alive at the time of sorting were collected. Fly heads were stored at room temperature in the dark (Lehane and Mail, 1985)

by covering samples with foil. They were placed within airtight bags containing silica beads so they would remain dry.

Pteridine analysis was performed at LSTM. The method used was adapted from Lehane & Mail (1985). Each fly head was homogenised in 3 mL of a 0.1 N NaOH buffer (adjusted to pH 10 with glycine (ca. 11.5 g/l)). The homogenate was then centrifuged until a pellet was formed, and the supernatant transferred into a glass cuvette. A fluorimeter (JENWAY®) was used to read raw fluorescence units (RFU) using filters of wavelength 340 nm for excitation and 450 nm for emission with the system gain set to 55%. A buffer reading and standard reading were taken were taken every hour to gain a standardised fluorescence to account for the fluctuations of the fluorimeter.

The heads belonging to the flies which had also been dissected for ovarian age were measured first, and a calibration curve plotted of ovarian age vs standardised fluorescence for both *G. pallidipes* and *G. swynnertoni*. The standardised fluorescence values of the flies of unknown age could then be used to calculate their estimated age using the regression equation y=a+bx where y = estimated age; a = x-axis intercept; b = slope of the regression line and x = standardised fluorescence. Generalised linear models for a Poisson distribution were then used to compare the frequency of flies with area and age as covariates.

2.2.5 Estimating fat content

A subsample of male flies collected during the October 2016 survey were examined for fat content. Only male flies were used as the presence of larvae in females, and their associated fat, will confound the results for the adult flies. Tsetse collected from Grumeti and North Ikorongo were pooled into 'Border' flies as few were caught alive in these sites. The head, legs and wings were removed from each fly and the size of the thorax measured. Flies were then placed in metal extraction trays with individual wells and dried in an oven at 70 °C overnight. The dry weight (DW) of each fly was then measured on the weighing scales (Hargrove, 1999). The extraction trays were then placed into a glass tray in the fume hood and the glass tray filled with chloroform until the metal tray was covered. The extraction trays have small holes in the top and bottom to allow the chloroform to enter each well.

The trays were immersed in chloroform for a total of 72 h to extract fat from the flies. The chloroform was drained and replaced with fresh chloroform every 24 h. After extracting fat, the trays were removed and placed into the oven at 70 °C overnight. Each whole fly was then reweighed for provide the residual dry weight (RDW), before being split into abdomen and thorax and the thorax weighed separately providing the thoracic residual dry weight

(TRDW). Fat content was calculated as the initial dry weight minus the post chloroform treatment weight (DW-RDW). The mean thoracic residual weight (TRDW) was then calculated for each species. Any flies which were 2 standard deviations lighter were discounted from the results as they are likely to be tenerals and not had chance to accumulate fat. An ANOVA was then conducted in R to test for significantly different mean fat contents and a Tukey HSD test applied if the initial ANOVA was significant.

2.3 Results

2.3.1 River transects – *G. pallidipes*

In sampling conducted in February 2016, the trap with the highest daily back transformed average trap catch of 288 was at the North Ikorongo site and the lowest daily back transformed average catch of 0.2 was at the Park Nyigoti site (Figure 9). When log transforming all catches had one added to them so that zero catches could be processed. For Band 7 surface reflectance, when all the transects are combined there is a significant relationship (P: 0.00453, adj. R^2: 0.157) where as Band 7 values increased, catches decreased (Figure 10). However, when examined by transect, only the Grumeti transect was significantly correlated (P: 0.0417, adj. R^2 0.317), with a similar slope to the overall relationship. For NDVI, when all the transects were combined there was a significant relationship (P: 0.0431, adj. R^2: 0.0723) where as NDVI values increased, average catches increased. By transect, none of the sites were significantly correlated. There was also a significant relationship when all sites were combined between average catches and land surface temperature (P: 0.00139, adj. R^2:0.200), where as temperatures increased, trap catches decreased. When examined by transect, only Robanda had a significant relationship (P: 0.0267, adj R^2: 0.375). Elevation was also significant with all sites combined (P: 0.0169, adj, R^2: 0.108), as elevation increased, catches increased. However, at North Ikorongo which had the only significant relationship for a site (P: 0.00225, adj. R^2: 0.627), the relationship was reversed – as elevation decreased, catches decreased.

Comparing the Hansen data tree cover image with average catch showed a significant relationship (P: 0.00843, adj. R^2: 0.134) for all the sites combined showing that as percentage tree cover increased, so did average catch (Figure 11). Again, significant relationships were not seen when examined by transect. The distance the trap was from a river was significantly related to the average catch with all sites combined (P: 0.00174, adj. R^2: 0.192), but only North Ikorongo had a significant relationship (P: 0.00611, adj. R^2: 0.539) by transect.

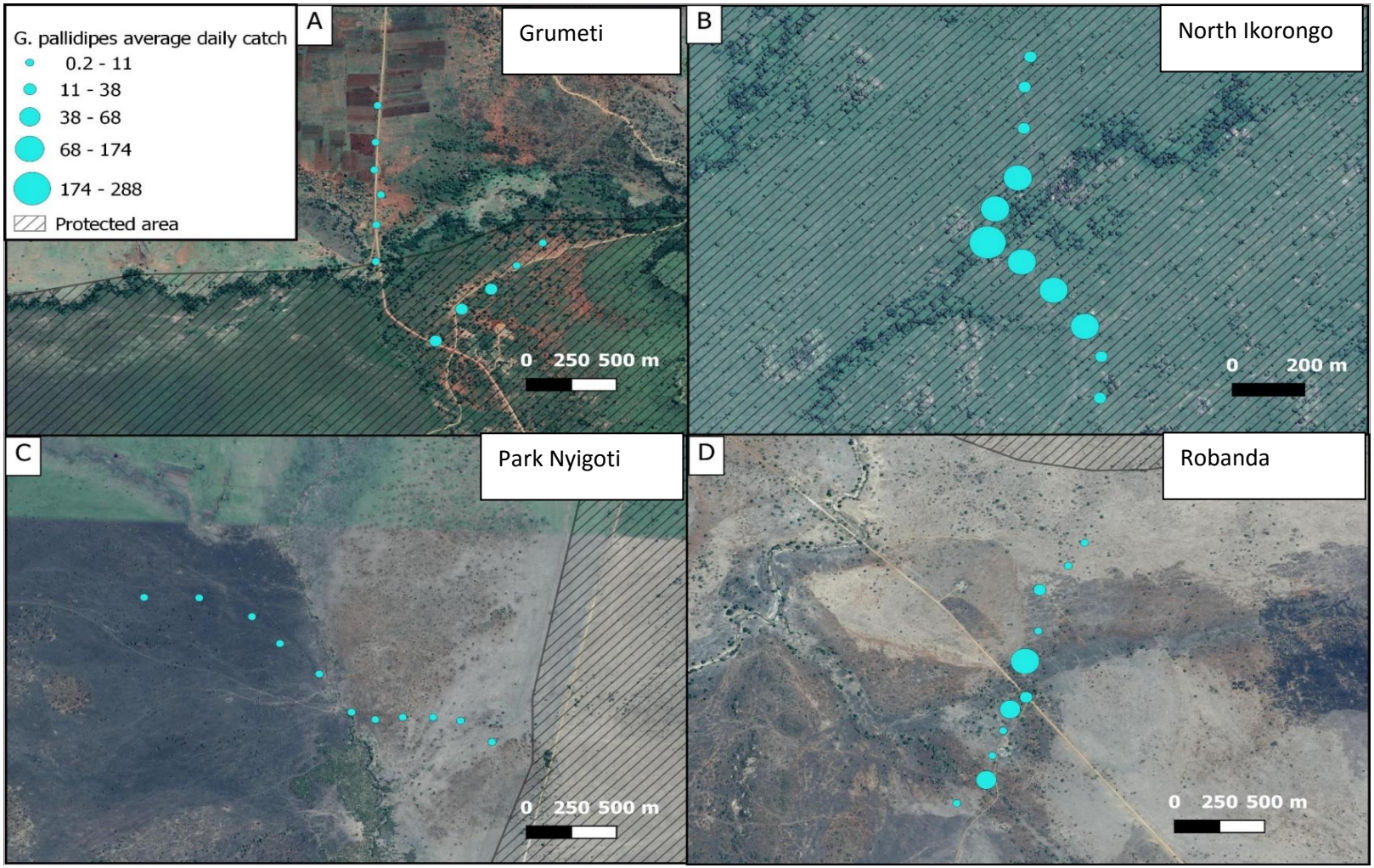

Figure 9. Average daily trap catches of G. pallidipes in the river transects at Grumeti (A), North Ikorongo (B), Park Nyigoti (C) and Robanda (D).

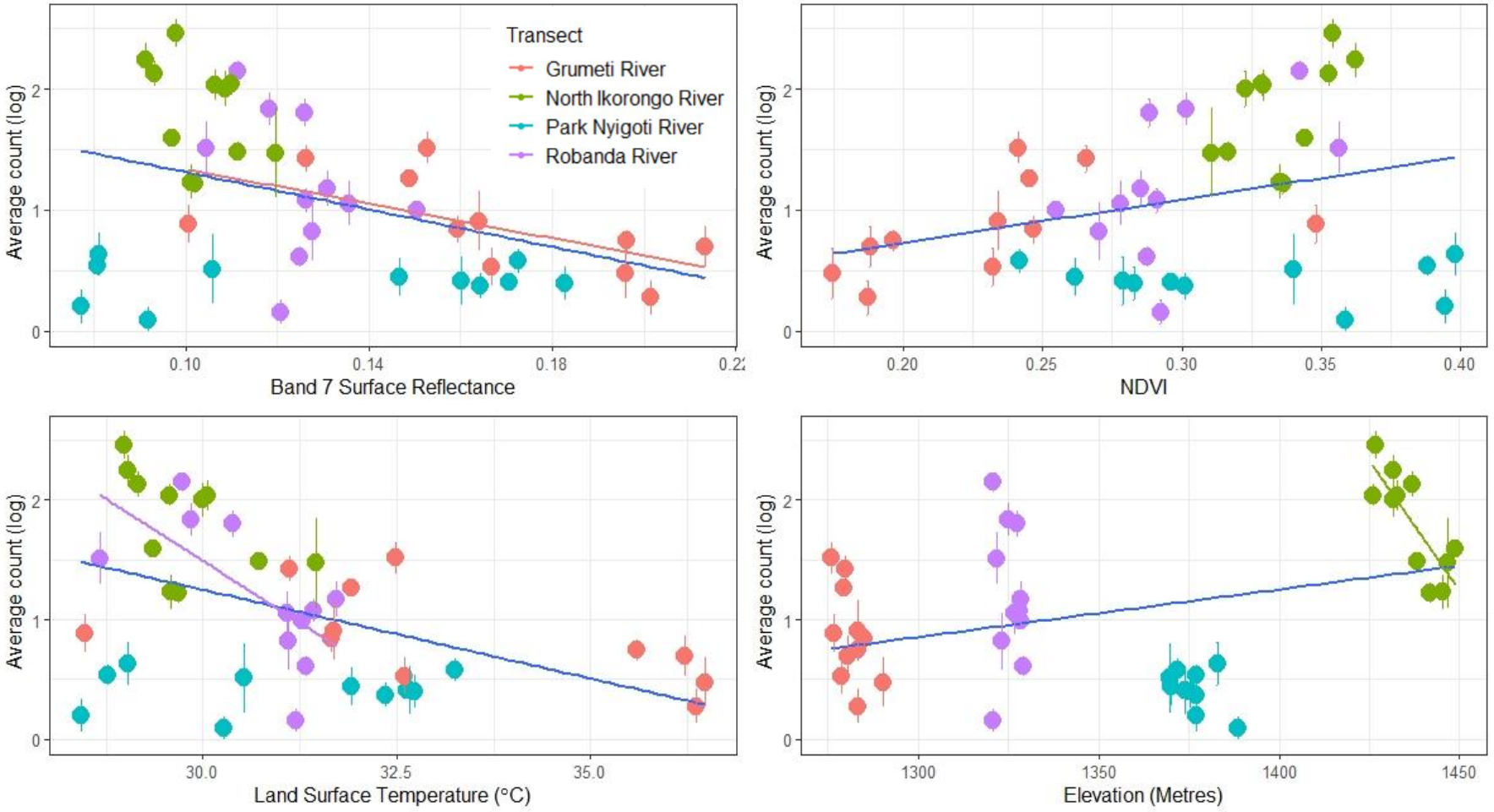

Figure 10. The log average G. pallidipes daily trap catch compared against Band 7 surface reflectance, NDVI, LST and elevation from sampling conducted in February 2016. A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. Error bars show standard error of the mean.

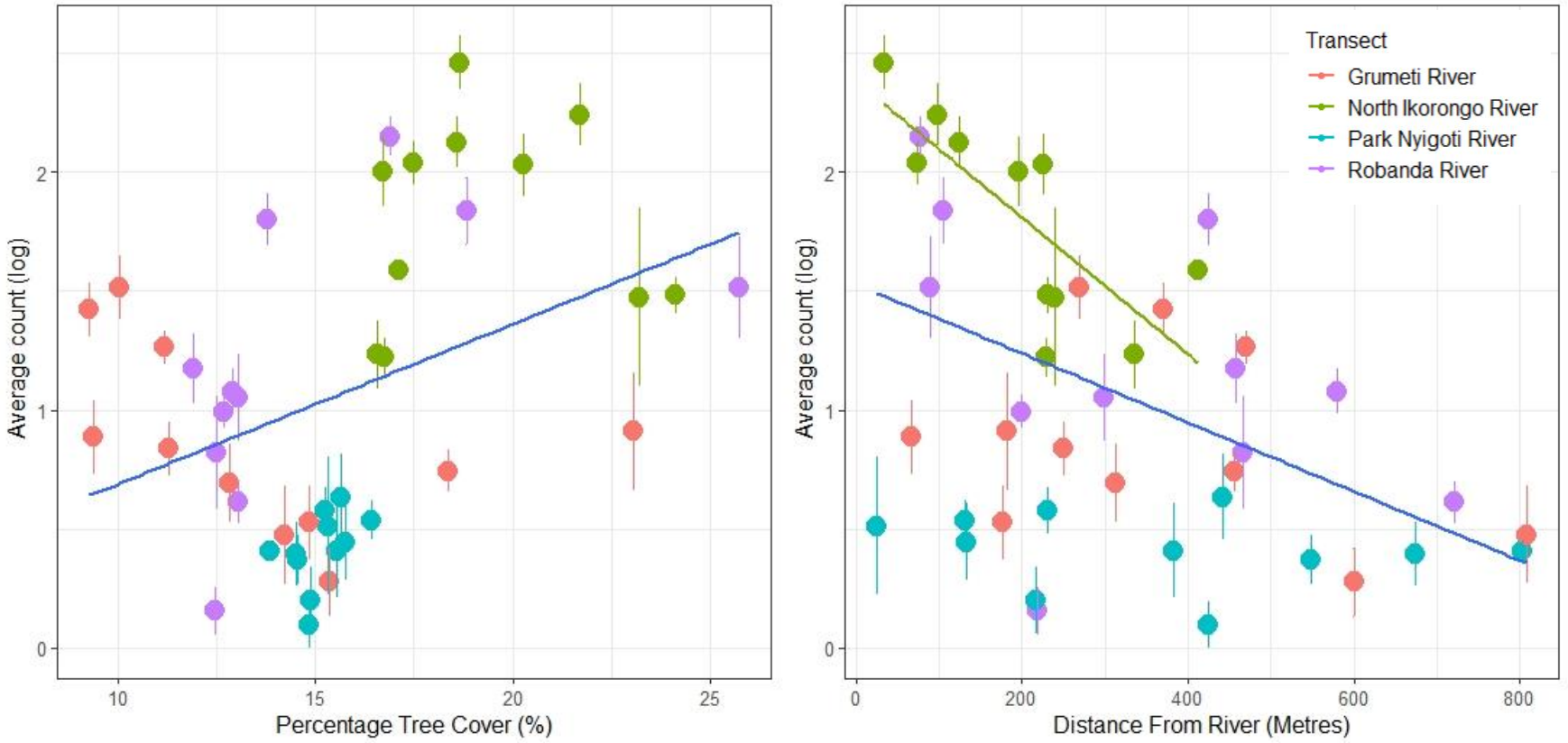

Figure 11. The log average G. pallidipes daily trap catch compared against Hansen percentage tree cover (L) and distance from a river (R). A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. This was from sampling conducted in February 2016. Error bars show standard error of the mean.

2.3.2 River transects – *G. swynnertoni*

In sampling conducted in February 2016, the trap with the highest back transformed daily average trap catch of 27 was at the North Ikorongo site and the lowest back transformed daily average catch of 0.3 was at the Park Nyigoti site (Figure 12). When log transforming all catches had one added to them so that zero catches could be processed.

There were significant relationships between average catch and Band 7 (P: 0.02, adj. R^2: 0.101) and average catch and land surface temperature (P: 0.0145, adj. R^2: 0.114) when all sites were combined (Figure 13). When all sites were combined there was no significant relationship between NDVI or elevation with average catch. When examined by transect, there was no significant relationships for any of Band 7, NDVI, land surface temperature or elevation. There were no significant relationships between percentage tree cover as determined by Hansen with average catch, or with distance from a river either combined or by individual transect (Figure 14).

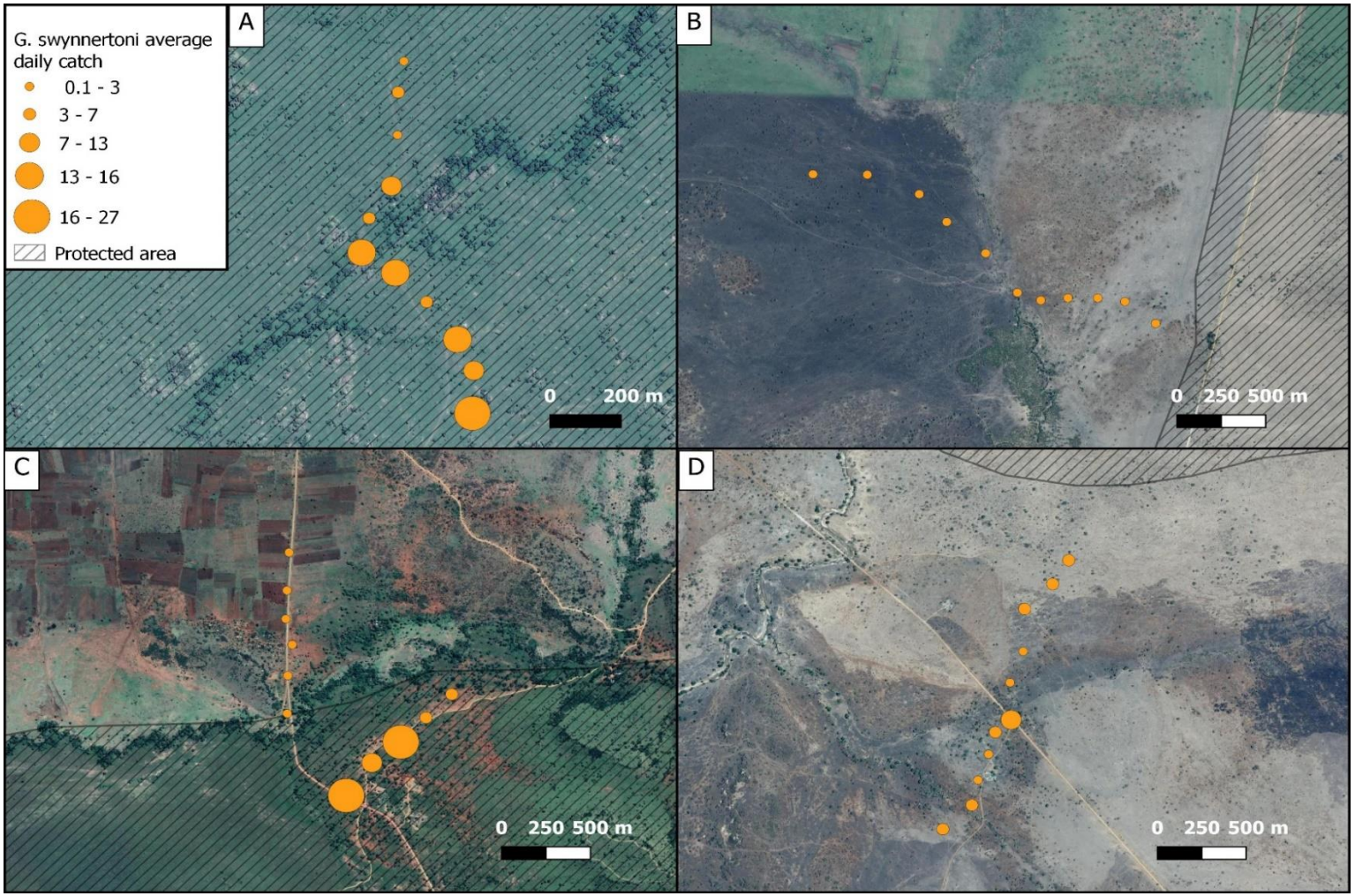

Figure 12. Average daily trap catches of G. swynnertoni on the river transects Grumeti (A), North Ikorongo (B), Park Nyigoti (C) and Robanda (D).

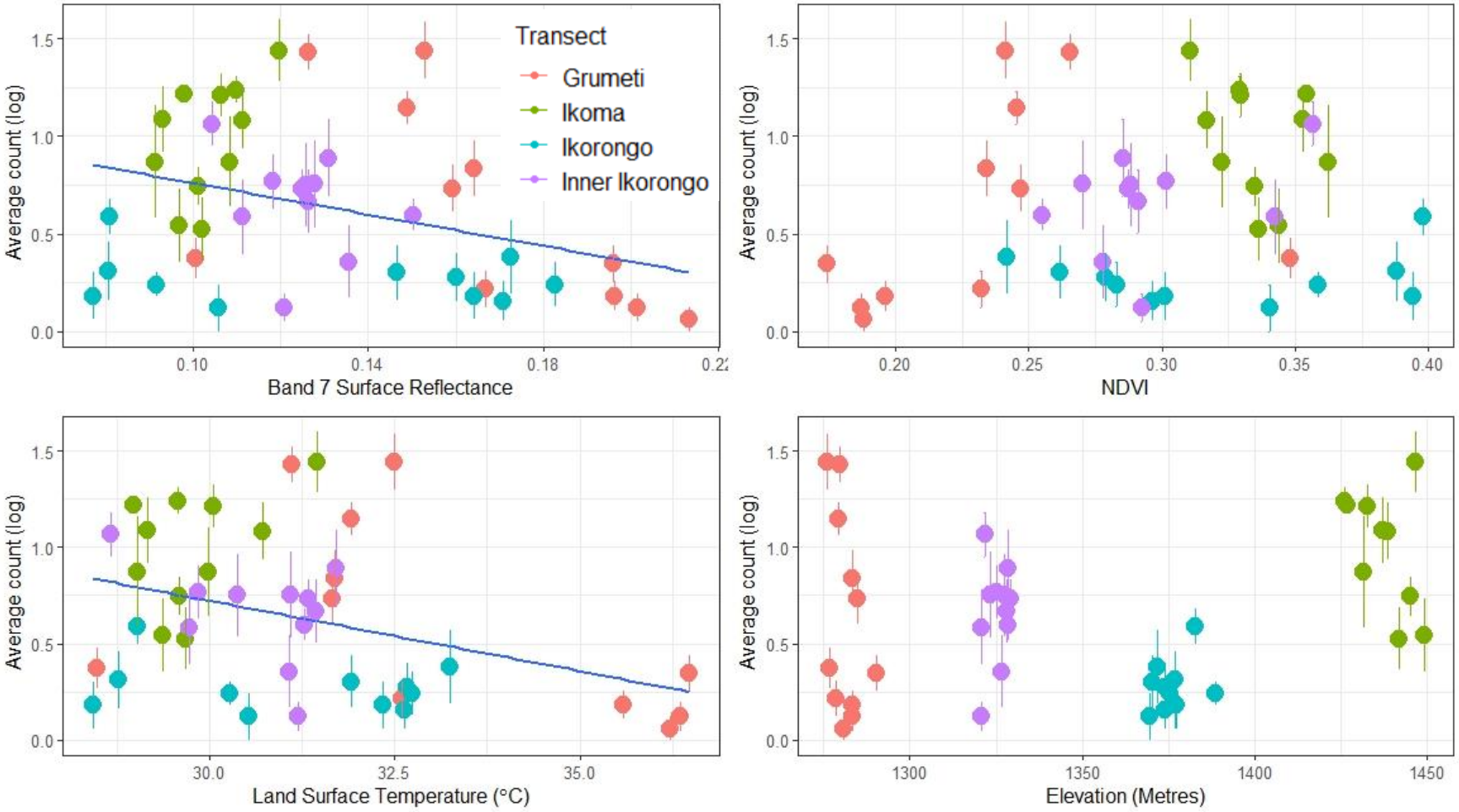

Figure 13. The log average G. swynnertoni daily trap catch compared against a manually calculated percentage tree cover (L) and distance from a river (R). A dark blue line indicates significant relationships with all sites combined. This was from sampling conducted in February 2016. Error bars show standard error of the mean.

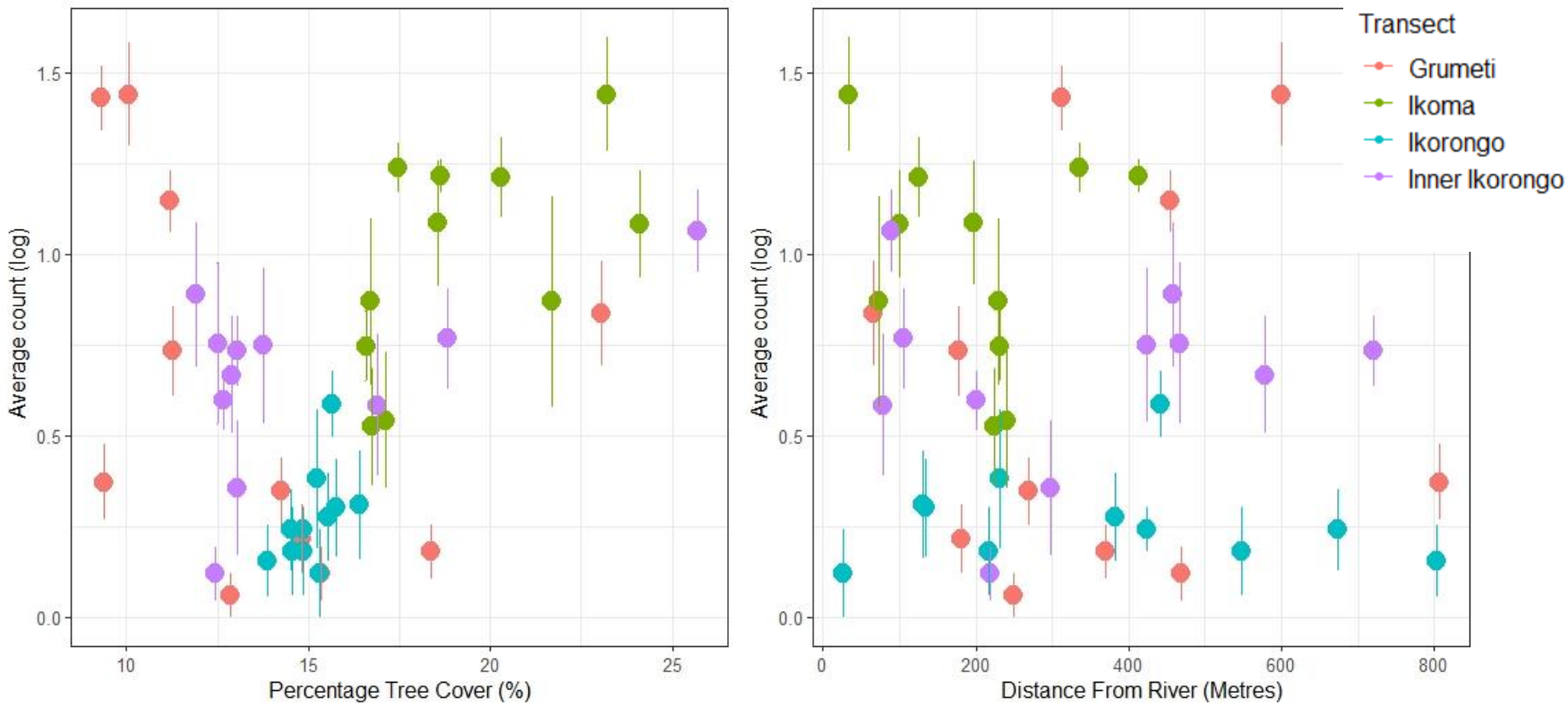

Figure 14. The log average G. swynnertoni daily trap catch compared against Hansen percentage tree cover (L) and distance from a river (R). This was from sampling conducted in February 2016. Error bars show standard error of the mean.

2.3.3 Grid sampling – *G. pallidipes*

In sampling conducted in October 2016, the trap with the highest back transformed daily average trap catch of 87 was at the North Ikorongo site and the lowest back transformed daily average catch of 0.3 was at the Inner Ikorongo site (Figure 15). When log transforming all catches had one added to them so that zero catches could be processed.

When all four grids were combined and compared with Band 7 surface reflectance, there was a significant negative relationship between catch and Band 7 (P: 2.48^{-5}, adj. R^2: 0.246 with catch decreasing as Band 7 increased (Figure 16). When examined by site, there was only a significant relationship at Ikoma (P: 0.00267, adj. R^2: 0.475), which had a very similar slope. It is possible that when examined by site and sample size is small, the effect at Ikoma which is the site furthest inside a protected area is stronger due to the lack of anthropogenic factors which may be influencing catch. For NDVI, there was no significant relationship found between average catches and NDVI when the four grids were combined however two of the sites did show significant relationships. Average catches at Ikoma (P: 0.0135, adj. R^2: 0.338) and inner Ikorongo (P: 0.00449, adj. R^2: 0.434) increased with NDVI. For LST, there was no significant relationship when all the sites were combined. By site, average catches at Ikoma did have a significant relationship (P: 0.000701, adj. R^2: 0.569), with catches decreasing as temperature increased. For elevation, there was a significant relationship with average catch when all sites were combined (P: 0.0206, adj. R^2: 0.0709), and two sites which were significant individually. Average catches at Grumeti were significant with elevation (P: 0.000109, adj. R^2: 0.645) where as elevation increased, catches increased. Average catches at North Ikorongo were also significant (P: 0.00799, adj. R^2: 0.363), but as elevation increased, catches decreased.

For grids, tree cover was examined using Hansen as done for the river transects, but also with a manually calculated percentage tree cover (Figure 17). For the manually calculated tree cover, with all sites combined this was significant with average catches (P: 2.92^{-5}, adj. R^2: 0.242). This was also significant for three of the sites individually. Ikoma had the most significant relationship (P: 0.000277, adj. R^2: 0.624), followed by inner Ikorongo (P: 0.00103, adj. R^2: 0.543), then North Ikorongo (P: 0.00258, adj. R^2: 0.452). Average catches at Grumeti were not significant with manually calculated tree cover. For Hansen, the significant sites were the same. North Ikorongo has the most significant result (P: 0.00858, adj. R^2: 0.357), followed by inner Ikorongo (P: 0.0373, adj. R^2: 0.238) and Ikoma (P: 0.0373, adj. R^2: 0.238). When sites were combined it was still significant (P: 0.00364, adj. R^2: 0.118). Hansen values were compared with the manually calculated percentage tree cover to see if the methods

were giving similar estimations and a significant relationship was found (P: 4.7^{-11}, adj. R^2: 0.509), where a higher Hansen value equalled a higher manually calculated value. Hansen was also found to be significantly correlated with Band 7 (P: 1.8^{-6}, adj. R^2: 0.190) and NDVI (P: 0.00308, adj. R^2: 0.0723) (Figure 18).

Distance from a river was not significantly related to average catch when all sites were combined. Two of the transects did have significant relationships but in contrasting ways. At North Ikorongo, as distance from a river increased, average catches decreased (P: 0.0204, adj. R^2: 0.280), whereas at Grumeti catches increased as distance from a river increased (P: 0.00106, adj. R^2: 0.515).

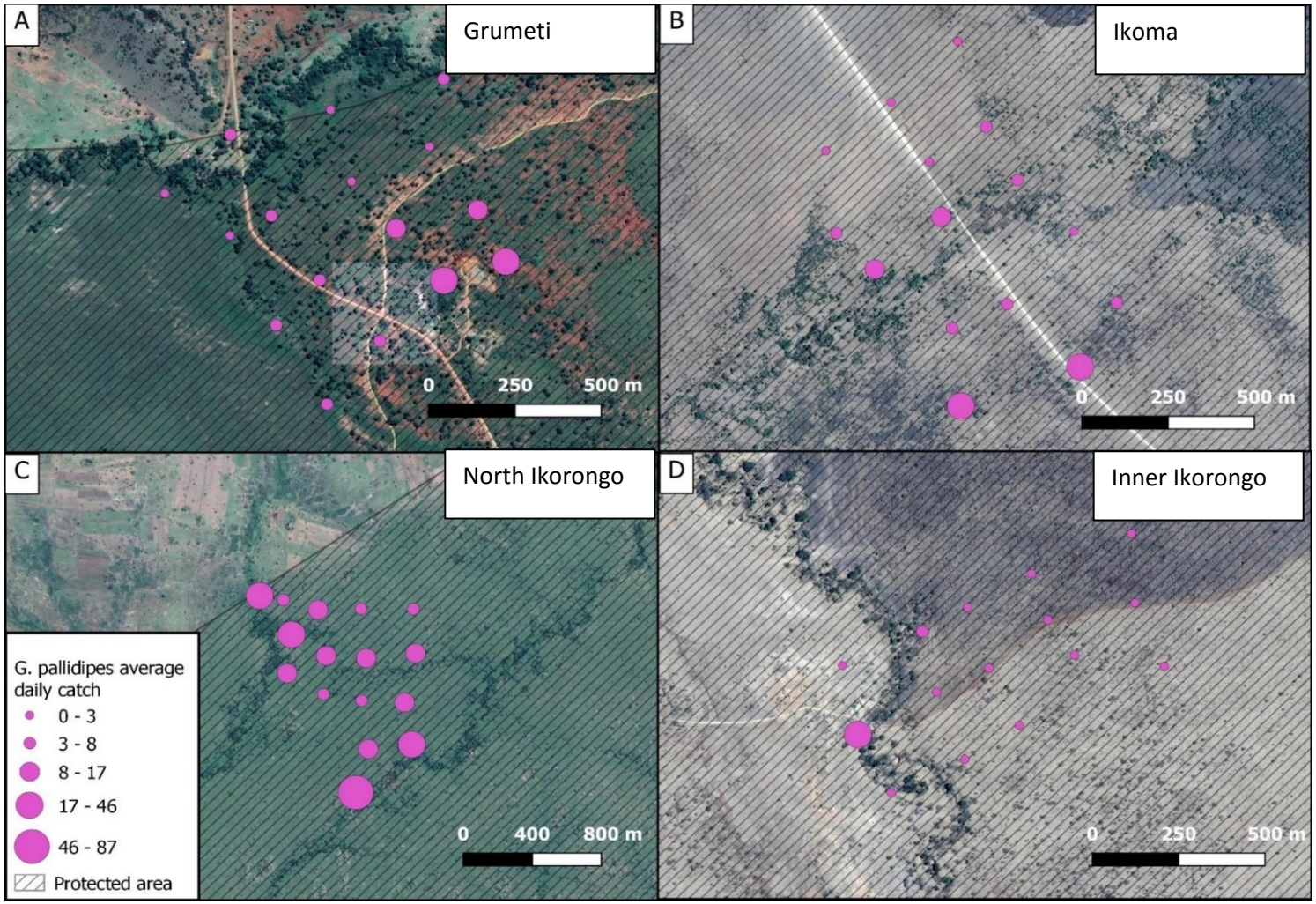

Figure 15. Average daily trap catches of G. pallidipes during the grid sampling at Grumeti (A), Ikoma (B), North Ikorongo (C) and inner Ikorongo (D).

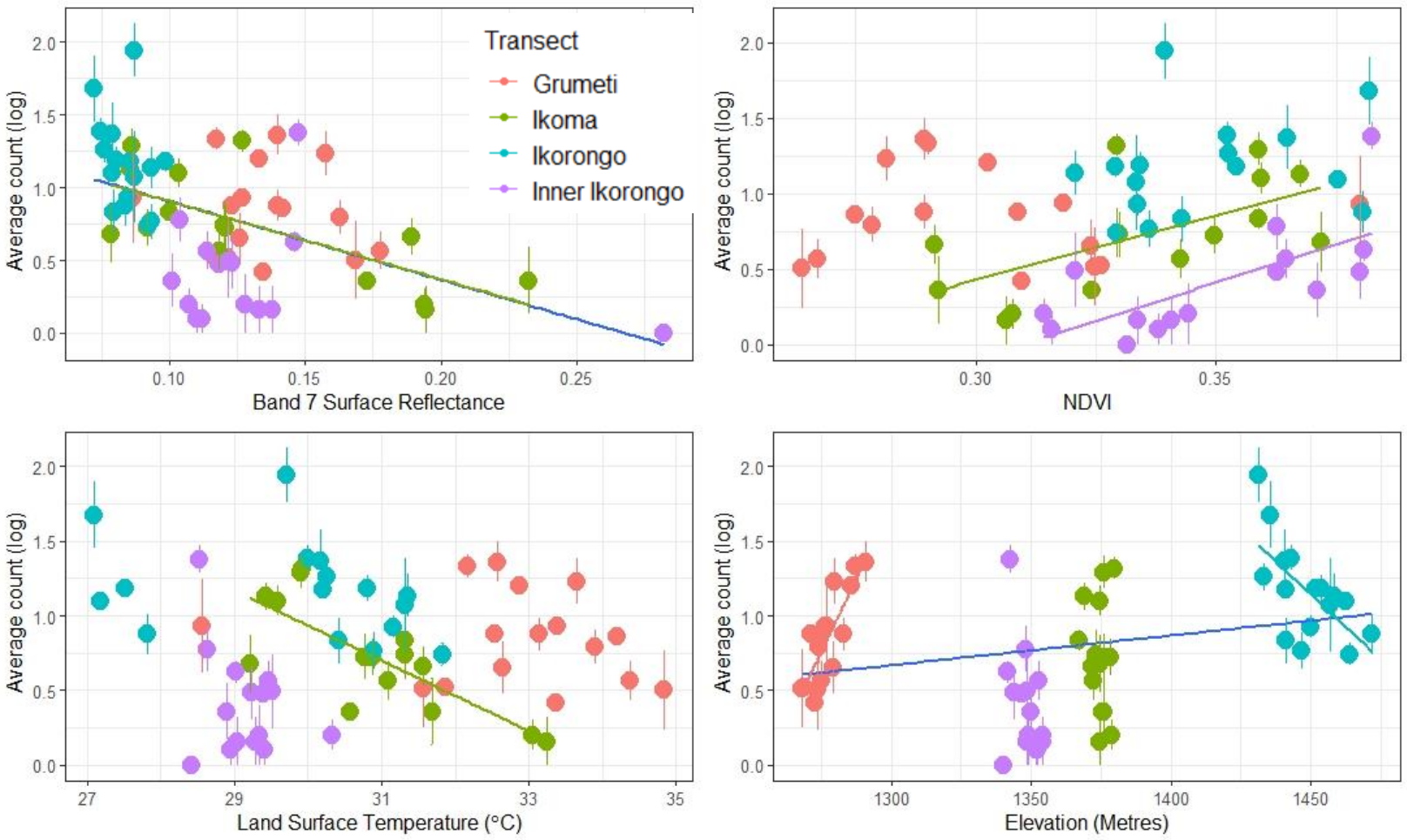

Figure 16. The log average G. pallidipes daily trap catch compared against Band 7 surface reflectance, NDVI, LST and elevation from sampling conducted in October 2016. A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. Error bars show standard error of the mean.

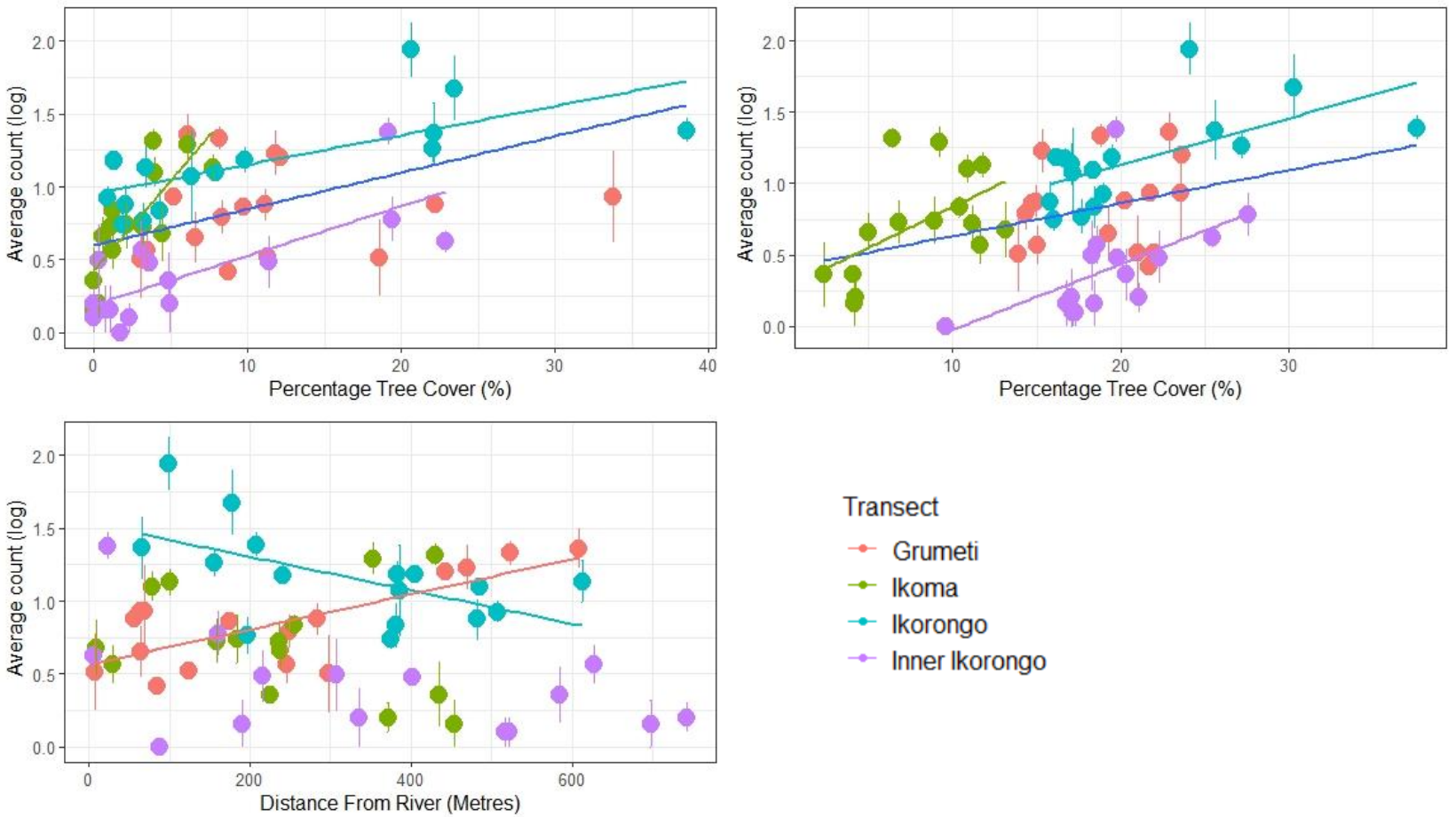

Figure 17. The log average G. pallidipes daily trap catch compared against a manually calculated percentage tree cover (TL), Hansen tree cover (TR) and distance from a river (LL). A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. This was from sampling conducted in October 2016. Error bars show standard error of the mean.

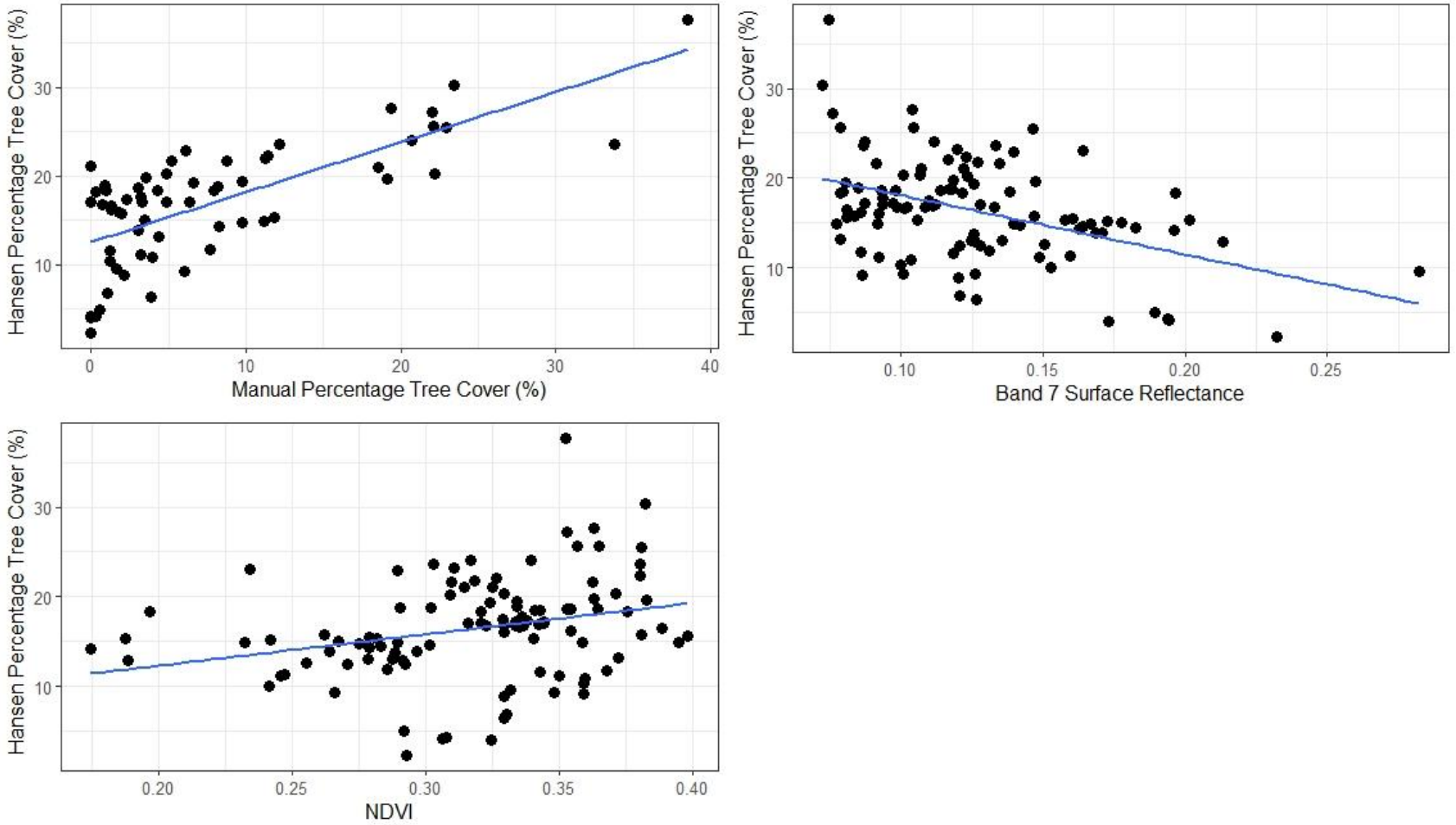

Figure 18. Hansen tree cover compared with a manually calculated percentage tree cover (TL), Band 7 surface reflectance and NDVI. Band 7 Surface Reflectance and NDVI values were taken from an October 2016 satellite image.

2.3.4 Grid sampling – *G. swynnertoni*

In sampling conducted in October 2016, the trap with the highest back transformed daily average trap catch of 36 was at the Ikoma site and the lowest back transformed daily average catch of 0 was at the North Ikorongo site (Figure 19). When log transforming all catches had one added to them so that zero catches could be processed.

For Band 7, NDVI, LST and elevation there were no significant relationships with average catch when all sites were combined (Figure 20). For Band 7, NDVI and land surface temperature there were also no significant relationships within sites. For elevation, Grumeti was the only site where a significant relationship was seen (P: 0.0111, adj. R^2: 0.335) where catches increased as elevation increased. This effect was not seen when sites were combined for elevation.

Percentage tree cover was significant with all sites combined for both the manually calculated method (P: 0.0199, adj. R^2: 0.0719) and with Hansen (P: 0.000137, adj. R^2: 0.204) (Figure 21). For both methods, as percentage tree cover increased, average catch decreased. Neither method for percentage tree cover showed a significant relationship by site. Average catch compared with distance from a river did not show any significant relationships either with all sites combined or individually.

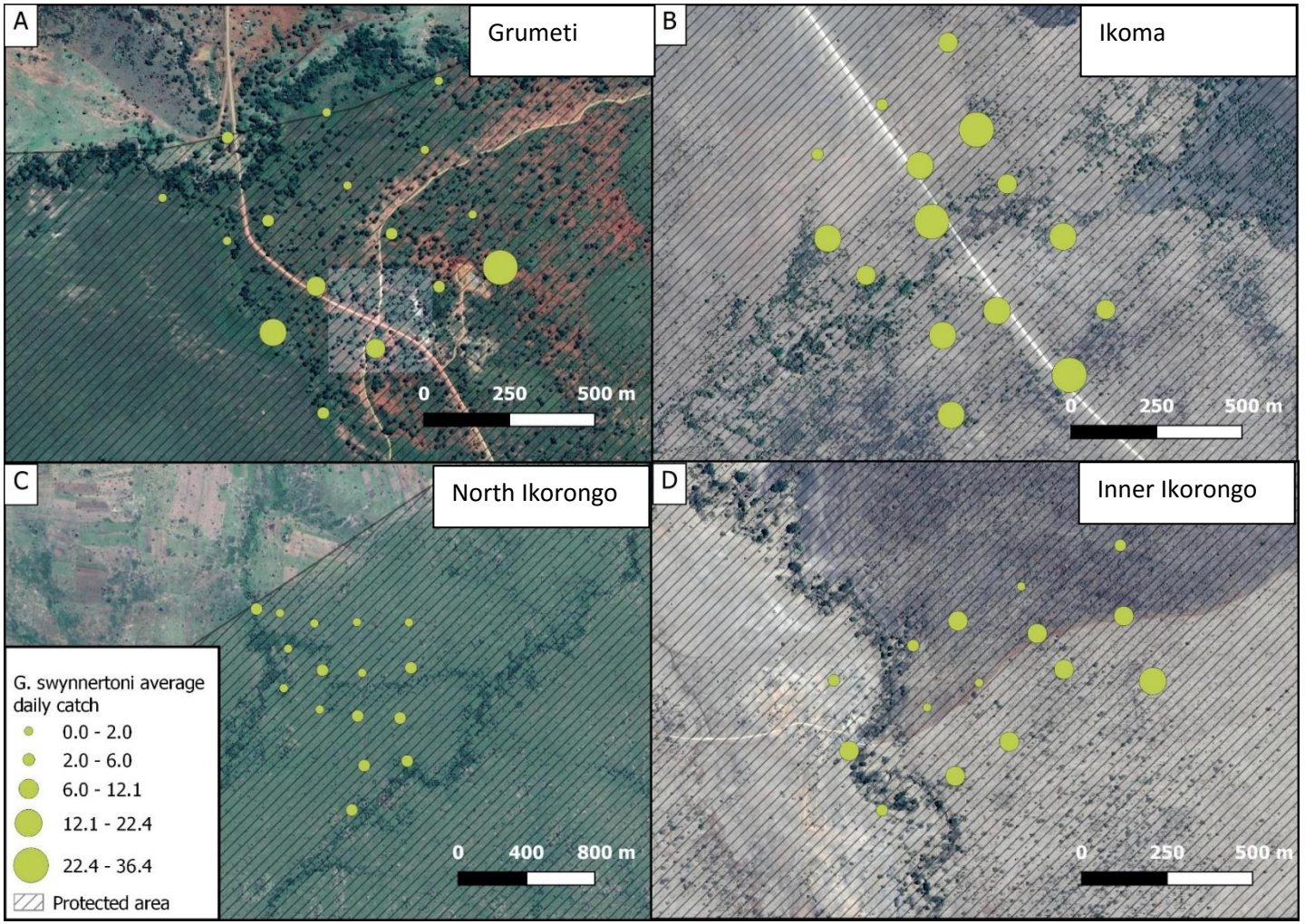

Figure 19. Average daily trap catches of G. swynnertoni during the grid sampling at Grumeti (A), Ikoma (B), North Ikorongo (C) and inner Ikorongo (D).

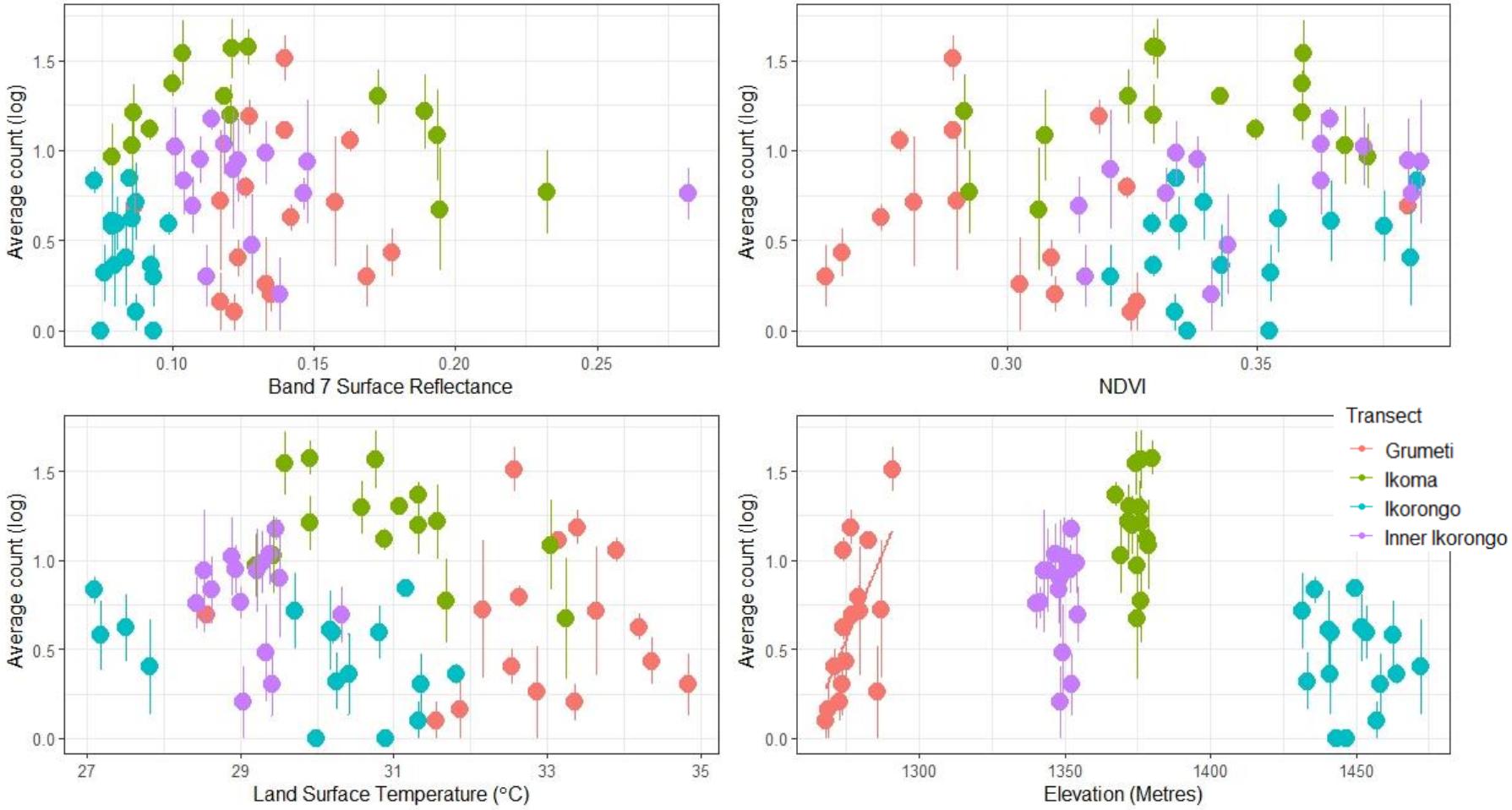

Figure 20. The log average G. swynnertoni daily trap catch compared against Band 7 surface reflectance, NDVI, LST and elevation from sampling conducted in October 2016. A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. Error bars show standard error of the mean.

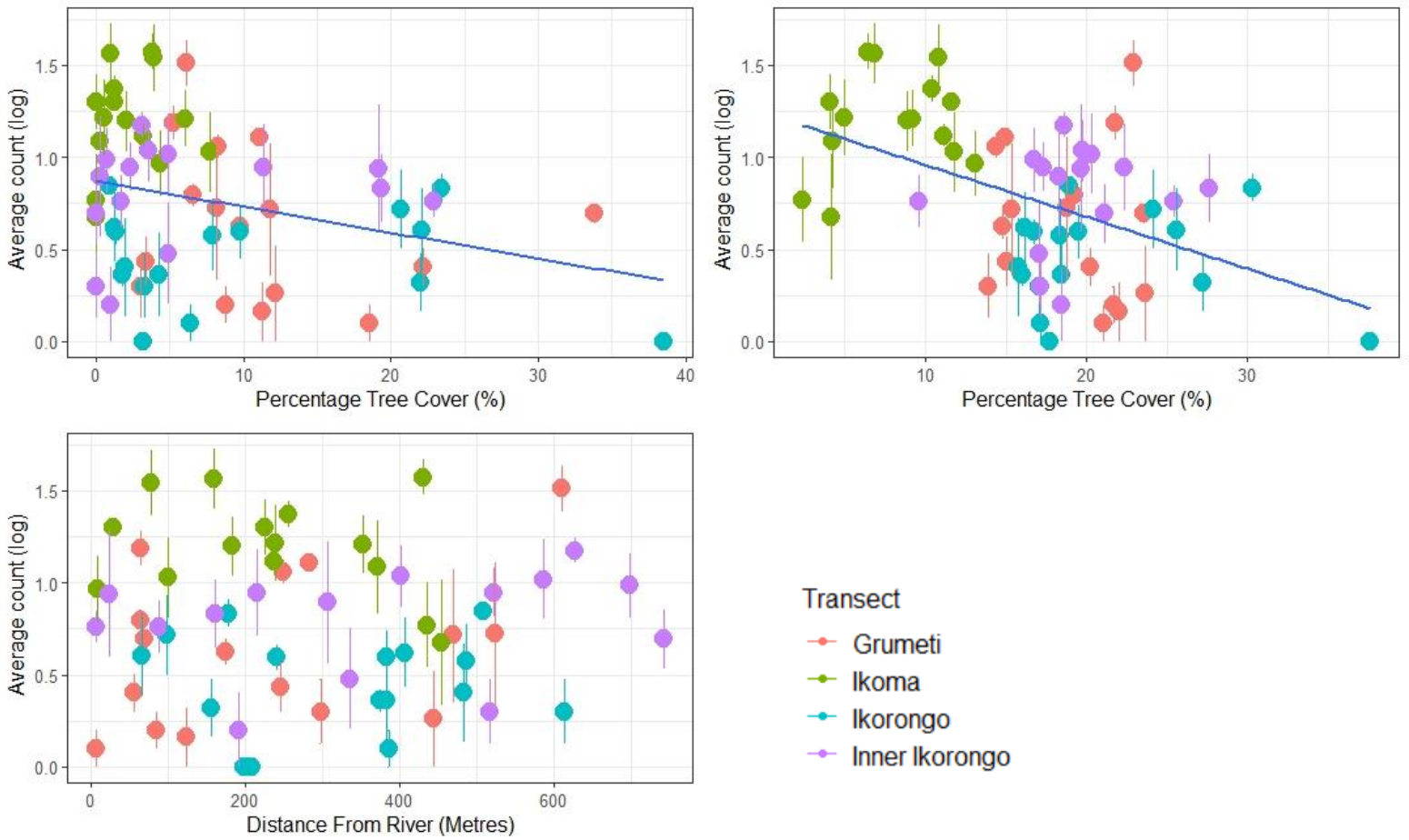

Figure 21. The log average G. swynnertoni daily trap catch compared against a manually calculated percentage tree cover (TL), Hansen tree cover (TR) and distance from a river (LL). A dark blue line indicates significant relationships with all sites combined. Lines matching the colour of specific sites indicate a significant relationship for that site. This was from sampling conducted in October 2016. Error bars show standard error of the mean

2.3.5 River transect and grid sampling combined – *G. pallidipes*

When all transects and grids are combined and the average catches compared with the environmental variables, almost all were significantly related (Figure 22). Band 7 surface reflectance was highly significant (P: 2.41^{-5}, adj. R^2: 0.150) and had the highest adj. R^2 value for the combined analysis. The overall trend showed that as Band 7 values increased, catches decreased, as they did when the river transects and grids were analysed within their time points. NDVI was the only variable which did not have a significant relationship with the combined catches. Land surface temperature analysis showed that as temperature increased between ~27 and ~38 °C, average catches decreased (P: 0.0134, adj. R^2: 0.0483). As elevation increased between ~1200 and ~1500 m, catches also significantly increased (P: 0.00252, adj. R^2: 0.0756).

As percentage tree cover as determined by Hansen increased, average catches also increased significantly (P: 0.00214, adj. R^2: 0.0783) (Figure 23). The distance from a river was also significantly related to average catch (P: 0.00346, adj. R^2: 0.070). Whilst significant relationships are reported in this section, many of them are weakly related with small adjusted R^2 values. The relationship between distance from a river and elevation was examined (Figure 24) to highlight how the two variables are geographically linked and that elevation is serving as a proxy for distance from river. The relationship was significant at six of the eight sites; at the Grumeti grid site (P: 0.000169, adj. R^2: 0.623), Grumeti river (P:0.0193, adj. R^2: 0.415), Ikorongo grid site (P: 0.000340, adj. R^2: 0.585), the inner Ikorongo grid (P: 2.74^{-6}, 0.813), the north Ikorongo river site (P:0.000494, adj. R^2: 0.730), Robanda (P: 0.0284, adj. R^2: 0.367) but not at Ikoma and Park Nyigoti, most likely due to the rivers in the shapefile in these two areas being very small and not well defined in the landscape.

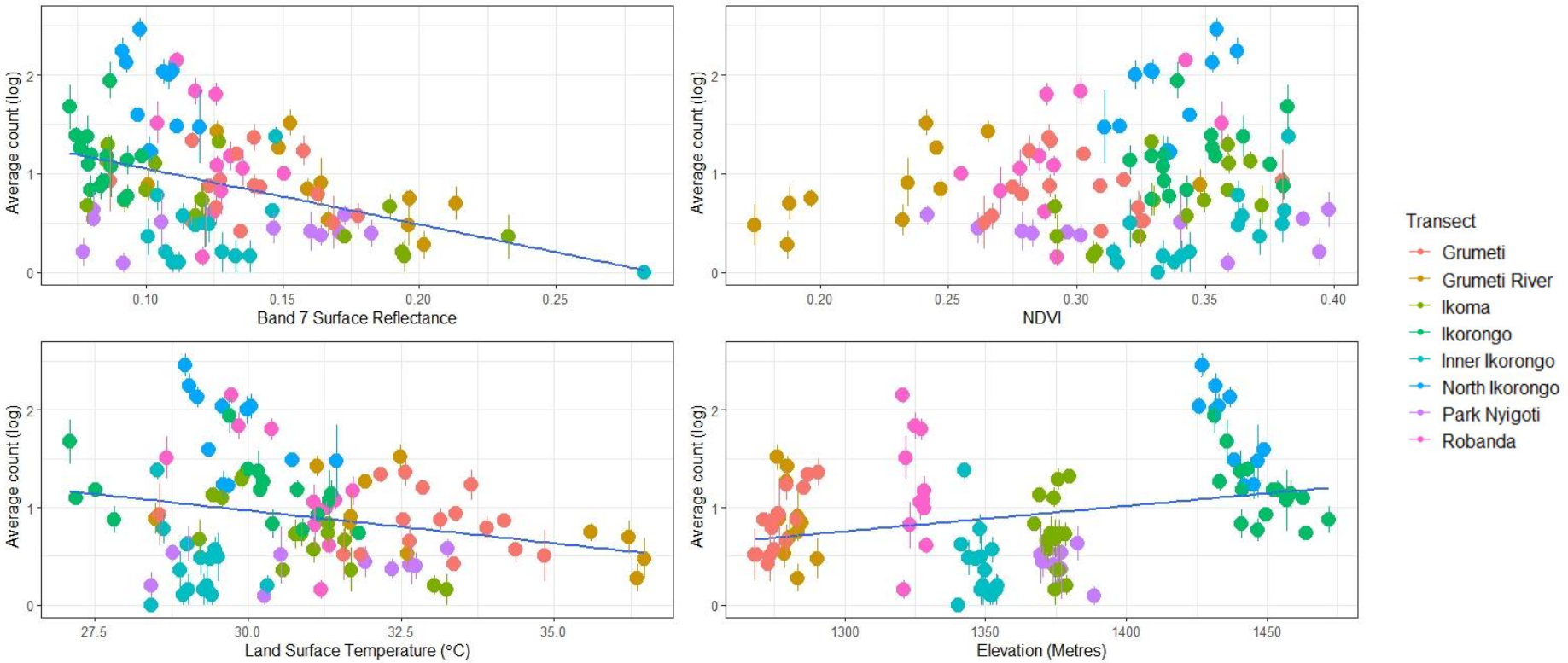

Figure 22. The log average G. pallidipes daily trap catch compared against Band 7 surface reflectance, NDVI, LST and elevation. A dark blue line indicates significant relationships with all sites combined. This combines data from the February 2016 and October 2016 sampling. Error bars show standard error of the mean

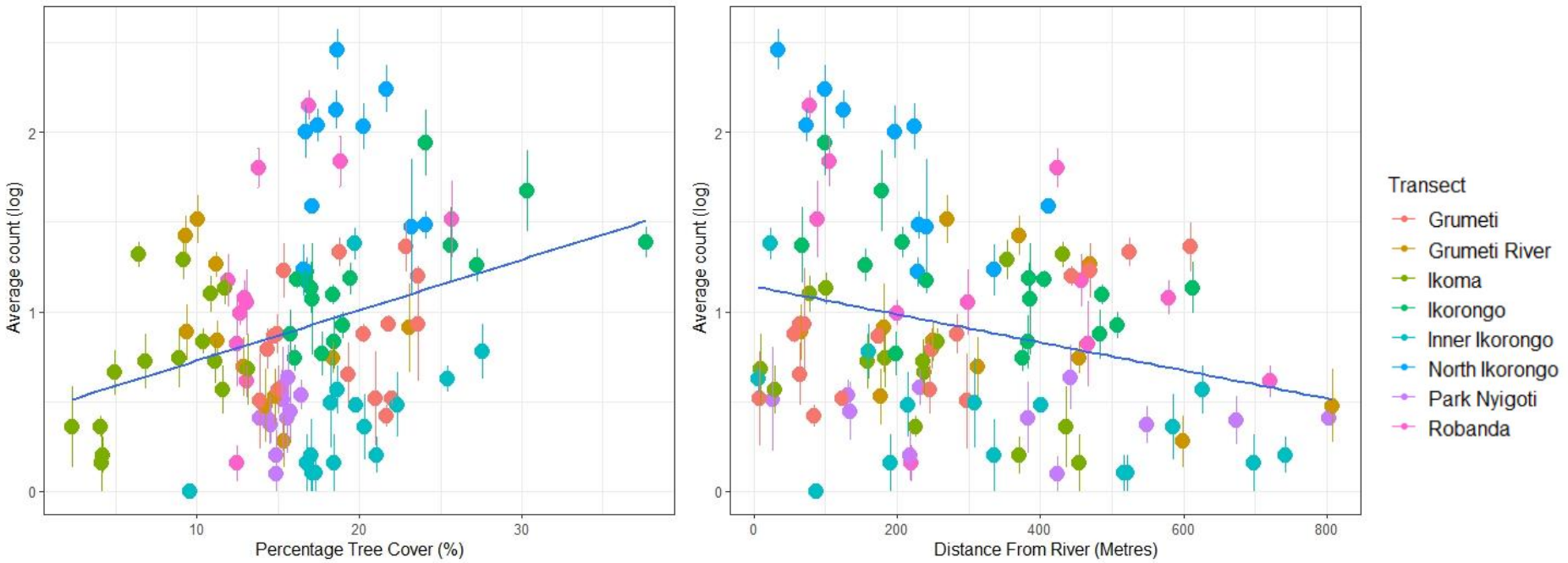

Figure 23. The log average G. pallidipes daily trap catch compared against percentage tree cover as determined by Hansen (L) and distance from a river (R). A dark blue line indicates significant relationships with all sites combined. This combines data from the February 2016 and October 2016 sampling. Error bars show standard error of the mean

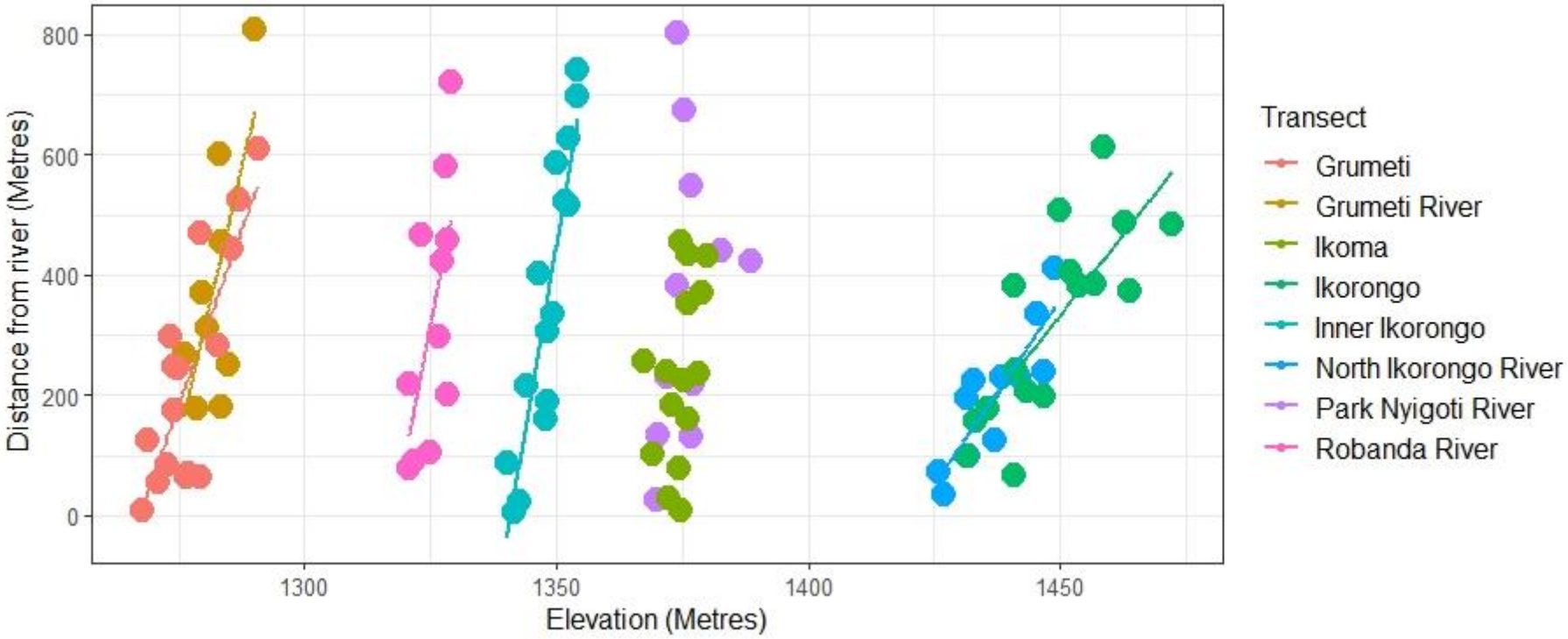

Figure 24. Showing the relationship between elevation and distance from river by transect. Coloured lines show significant relationships.

2.3.6 River transects and grid sampling combined – *G. swynnertoni*

Environmental variables were generally not significantly correlated with *G. swynnertoni* average catches when data from river transects and the grid sampling was combined. There were no significant relationships with Band 7, NDVI, elevation or distance from a river (Figure 25).

There was however a significant relationship between average catch and LST (P: 0.0240, adj. R^2: 0.0389) where as temperature increased, average catches decreased (Figure 26). The percentage tree cover as determined by Hansen was also significantly correlated with average catch (P: 0.0184, adj. R^2: 0.0431), where as percentage tree cover increased, average catches decreased.

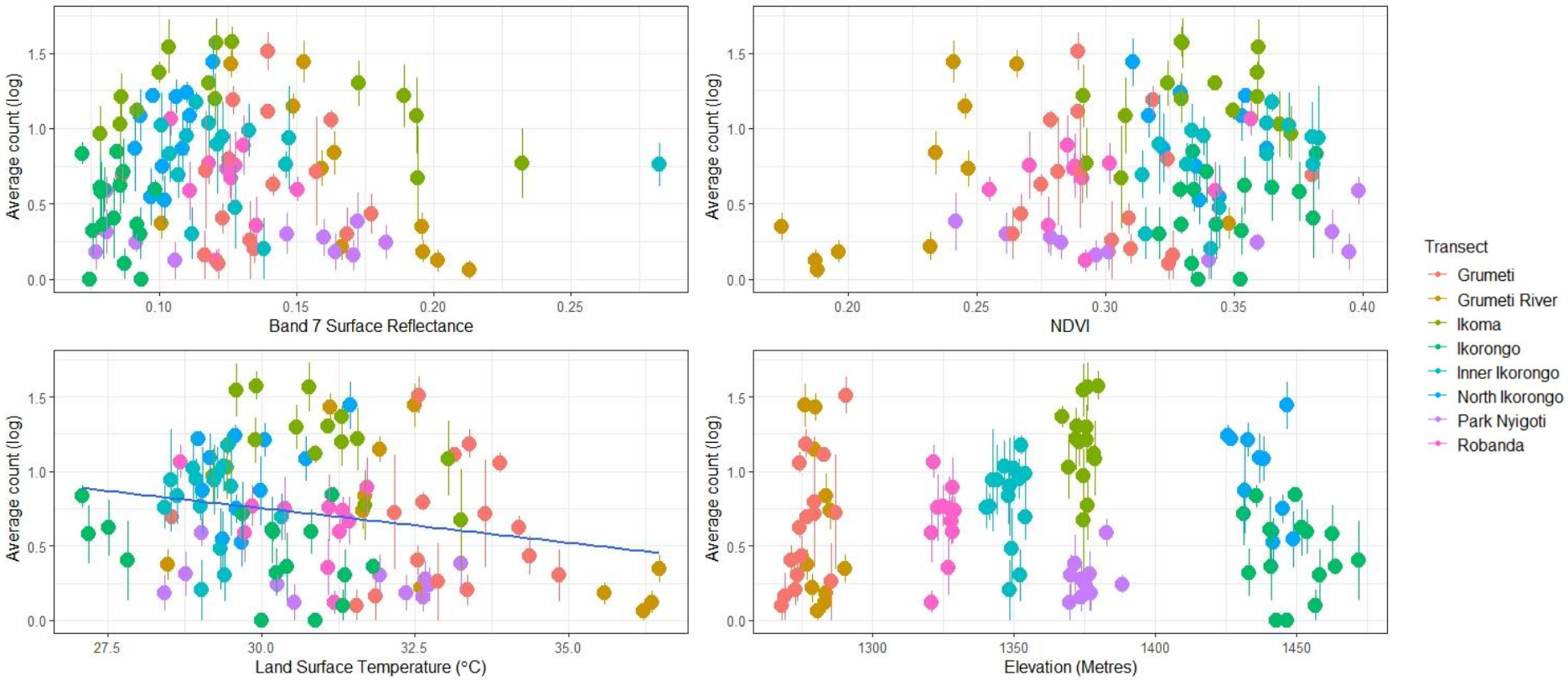

Figure 25. The log average G. swynnertoni daily trap catch compared against Band 7 surface reflectance, NDVI, LST and elevation. A dark blue line indicates significant relationships with all sites combined. This combines data from the February 2016 and October 2016 sampling. Error bars show standard error of the mean

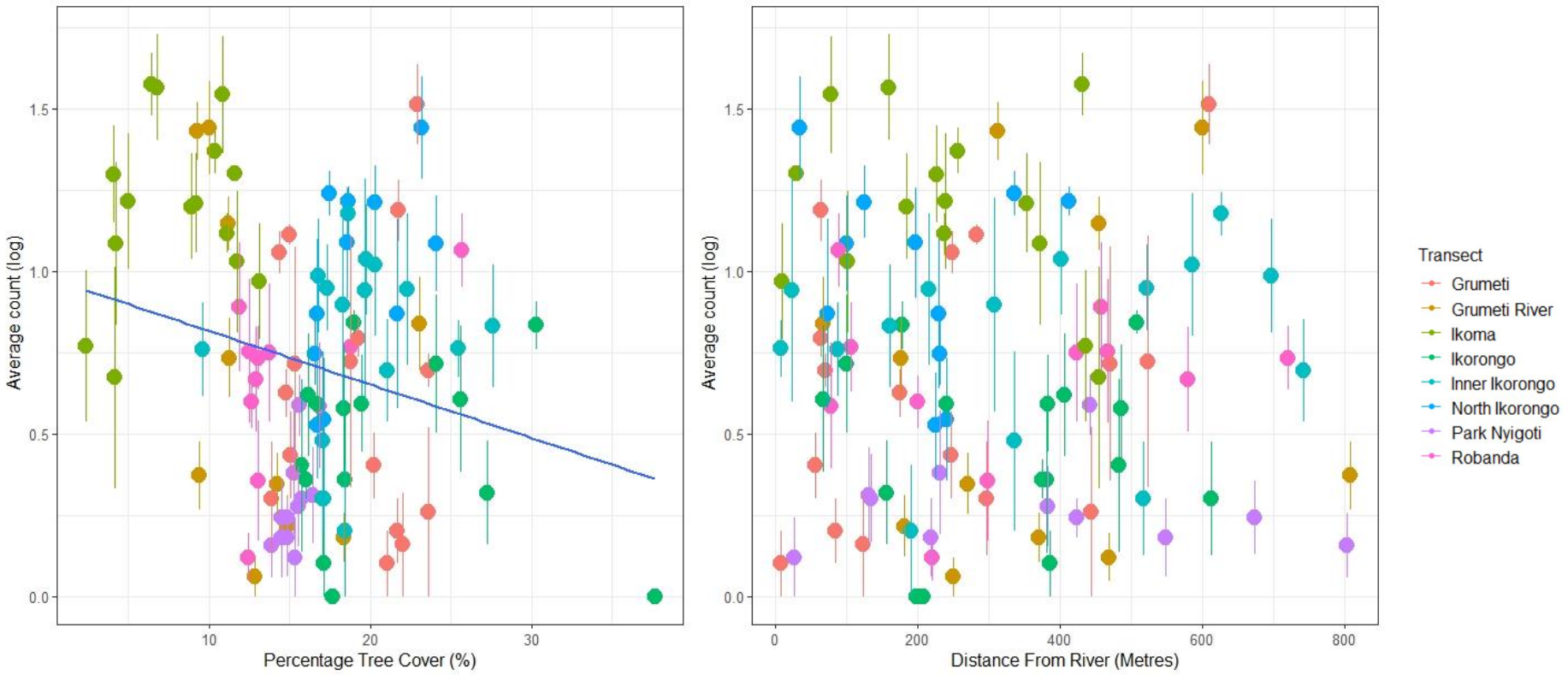

Figure 26. The log average G. swynnertoni daily trap catch compared against a percentage tree cover as determined by Hansen (L) and distance from a river (R). A dark blue line indicates significant relationships with all sites combined. This combines data from the February 2016 and October 2016 sampling. Error bars show standard error of the mean

2.3.7 Estimating the age of tsetse

Female flies collected from each grid set during October 2016 were analysed to determine their age with the pteridine method. To calibrate pteridine levels for the species, a subsample (~60) were dissected to determine their ovarian age category. The head of the dissected fly was kept and measured for the amount of pteridine with a fluorimeter. After calculating the standardised fluorescence values, they were compared with the ovarian age category using a linear regression model.

For calibration of *G. pallidipes*, the relationship between ovarian age and standardised fluorescence was significant (P: 2.36^{-10}, adj. R^2: 0.476) with fluorescence increasing as ovarian age increased (Figure 27). Variation of standardised fluorescence within ovarian age categories appeared to generally increase in flies over 30 days old.

For *G. swynnertoni*, the relationship between ovarian age and standardised fluorescence was also significant (P: 7.83^{-16}, adj. R^2: 0.646), with the same trend that was seen for *G. pallidipes* (Figure 27). There appeared to be less variation in standardised fluorescence values in flies over 30 days old than there was for *G. pallidipes* as reflected in the higher adj. R^2 value.

After the remaining fly heads had been processed for pteridine only and an estimated age calculated using the equation of the line, the frequency of flies in each age category was compared across areas. For *G. pallidipes*, the majority of flies in each area were young and the oldest flies were found in the smallest numbers (Figure 28). The oldest fly was found at North Ikorongo, with no flies older than 80 days old found at Inner Ikorongo or Ikoma. The decline in frequency of flies in each age group as they get older appears less steep at Grumeti. A generalised linear model for a Poisson distribution found significant differences between Ikoma and Grumeti (P: 0.0154), but not between any of the other sites.

For *G. swynnertoni*, sample sizes were only large enough for analysis at Ikoma, Inner Ikorongo and Grumeti. At three sites it appears as though the proportion of flies aged 0-10 days old was low, then rose to a peak between 11 - 40 days old, before decreasing as flies got older (Figure 29). The oldest flies were found at Ikoma and Inner Ikorongo. At Grumeti, there was a steeper drop between the category with the highest frequency (age 21-30 d old) and the categories next oldest or youngest. A generalised linear model for a poisson distribution found no significant differences across sites.

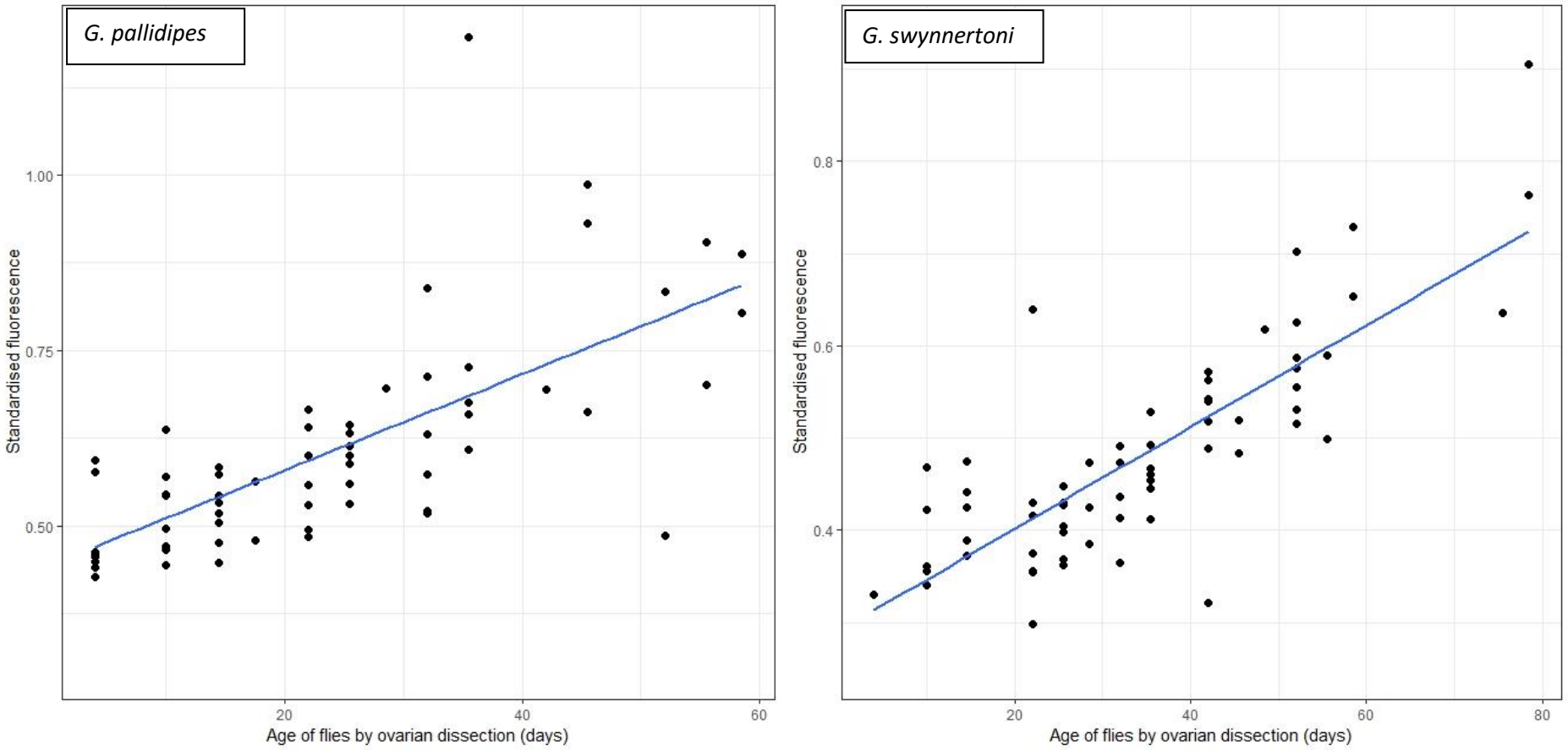

Figure 27. Calibration curves for pteridine assay. The age of flies by ovarian dissection was compared to the standardised fluorescence value from the pteridine to give a straight line with equations y=0.00685 x + 0.4422 and y=0.005516 x + 0.2913 for G. pallidipes (L, n=63) and G. swynnertoni (R, n=64) respectively.

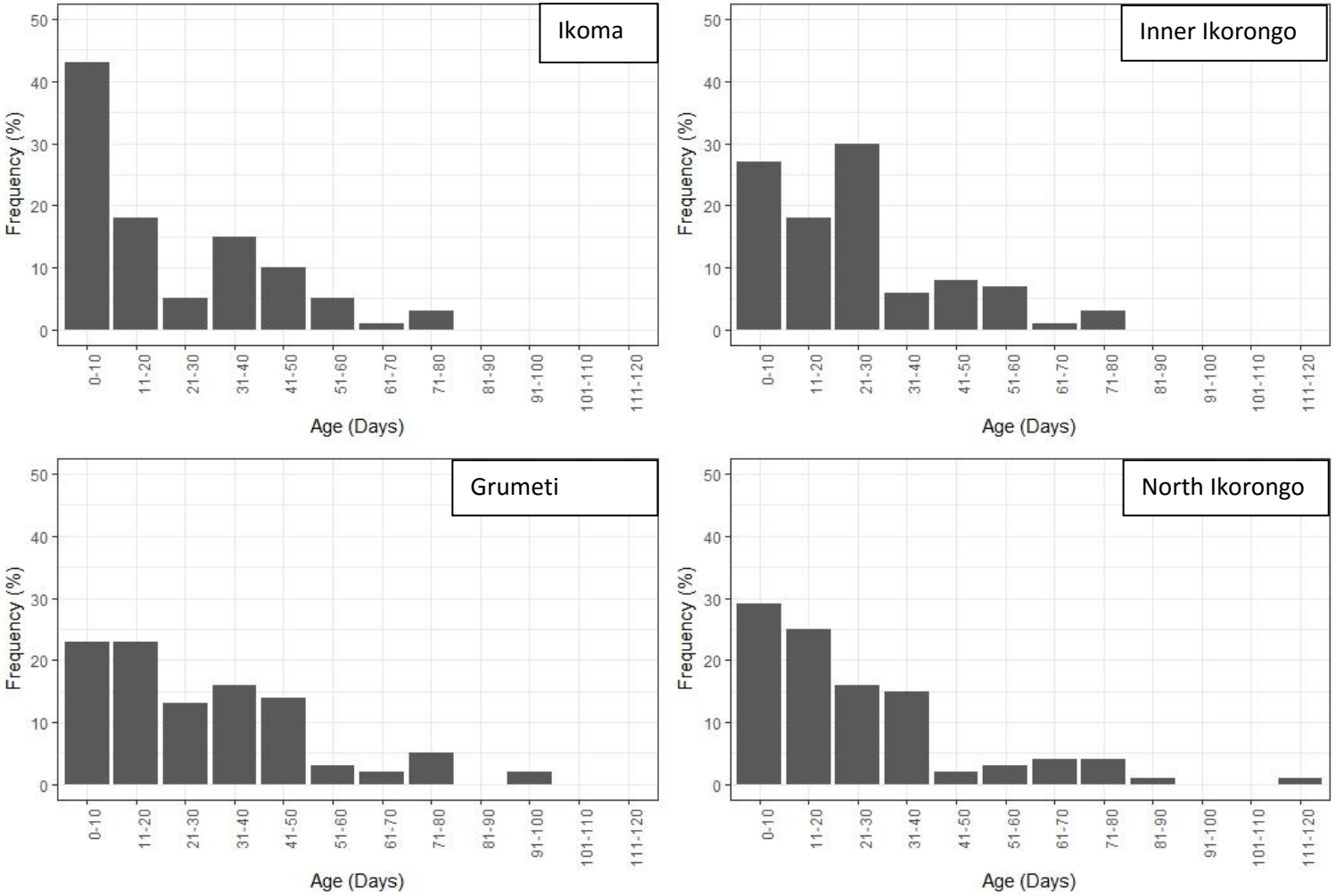

Figure 28. Composition of age groups of G. pallidipes population at Ikoma (TL, n=130), Inner Ikorongo (TR, n=97), Grumeti (LL, n=143) and North Ikorongo (LR, n=135).

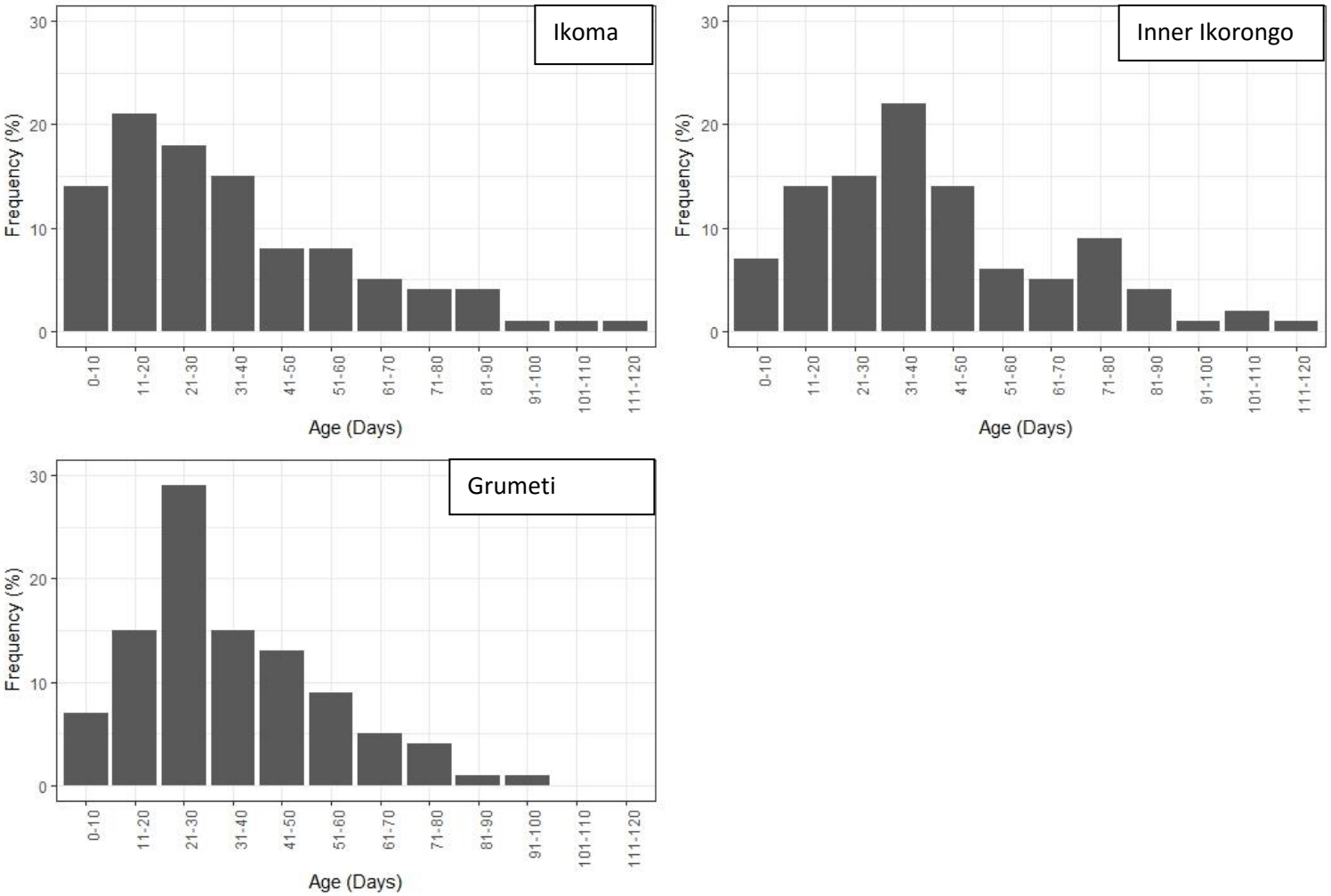

Figure 29. Composition of age groups of G. swynnertoni population at Ikoma (TL, n=208), Inner Ikorongo (TR, n=203) and Grumeti (LL, n=170).

2.3.8 Estimating fat content

With all *G. pallidipes* flies included (n=97) the mean dried thoracic weight was 5.34 ± 1.03 mg. After flies with a thoracic weight of less than or greater than 2SD were removed, this left a sample size of 90. Three of the removed flies were from Inner Ikorongo, two from the border and two from the Serengeti National Park group. The new mean dried thoracic weight for all *G. pallidipes* was 5.51 ± 0.86 mg. The mean fat content for each group of flies was calculated (Table 1) and the group of flies with the largest mean fat content was the group containing flies from the border areas.

Table 1. Mean fat content and SD for each group of G. pallidipes flies collected in October 2016.

Location	N	Mean fat content (mg)	SD
Border	53	1.426	0.939
SNP	17	0.794	0.649
Inner Ikorongo	20	0.860	0.804

A one-way ANOVA was conducted on the three groups and gave a significant result (P: 0.00718). A Tukey multiple-pairwise comparison was then conducted on the ANOVA result which gave a significant difference between the means at the border and the Serengeti National Park (P: 0.0272), and a significant difference between the border and Inner Ikorongo (P: 0.0377). There was no significant difference between the means from the Serengeti National Park and Inner Ikorongo (P: 0.971).

With all *G. swynnertoni* flies included (n=243) the mean dried thoracic weight was 3.90 ± 0.55 mg. After flies with a thoracic weight of less than or greater than 2SD were removed, this left a sample size of 236. All seven removed flies were all from the Serengeti National Park group. The new mean dried thoracic weight for all *G. swynnertoni* was 3.95 ± 0.47 mg. The mean fat content for each group of flies was calculated (Table 2) and the group of flies with the largest mean fat content was again the group containing flies from the border areas. A one-way ANOVA was conducted on the three groups and was not significant (P: 0.218).

Table 2. Mean fat content and SD for each group of G. swynnertoni flies collected in October 2016.

Location	N	Mean fat content (mg)	SD
Border	70	0.716	0.511
SNP	86	0.678	0.612
Inner Ikorongo	80	0.544	0.767

2.4 Discussion

2.4.1 River transects

The river transects in February 2016 found significant relationships between average *G. pallidipes* catch and every environmental variable when all four transects were combined in the same data set. The trend of these relationships all matched with results described in earlier literature (Kitron *et al.*, 1996; Sciarretta *et al.*, 2010; Lord *et al.*, 2018). Percentage tree cover and distance from river are very closely related as the water in the rivers will allow riparian vegetation to grow which provides suitable habitat for *G. pallidipes*. Catches also increase with NDVI, the measure of vegetation, which is also closely linked with rivers as riparian vegetation is normally much more dense than surrounding vegetation and will give higher values from satellite images. Band 7 is also generally higher in areas around rivers as the riparian vegetation will provide more shade so soil moisture is more likely to remain and these are areas where flies are less likely to desiccate. The shade provided will also reduce land surface temperature.

However, within each transect site, significant relationships were rarely seen. At Grumeti, only Band 7 was significantly related to average catch and at Robanda only LST was significant. Elevation at North Ikorongo was significant, however the trend was opposite to the overall trend and was likely due to the effect of the site as the transect extended either side of a river at the bottom of a valley, where there was a large amount of riparian vegetation. This is reinforced by the fact that distance from a river was also a significant factor for average catch at North Ikorongo. This shows that when enough data points are included over a large area, trends can be seen, however with small sample sizes over small areas, it is generally not possible to replicate the larger-scale results.

For *G. swynnertoni* average catches on the river transects, there were no significant relationships by transect however with all transects combined Band 7 and LST were significant. This likely highlights the limits of both tsetse species that they cannot survive in very hot, dry areas. The fact that there were no significant relationships with percentage tree cover and distance from river highlights the difference between the species, as *G. swynnertoni* are known to be more able to spread more equally across the habitat in an area (Lord *et al.*, 2018).

2.4.2 Grid sampling

The results from the October 2016 grid sampling showed slightly different results to those gained from the river transects for *G. pallidipes*. While Band 7, elevation and percentage tree cover were still significant with all grids combined; NDVI, LST and distance from a river were not. When all grids are combined, it appears that the data points from Grumeti fit less well with general patterns. Within grids, Grumeti was significantly correlated with distance from a river and elevation, however the trends were opposite to what is normally expected for *G. pallidipes*. When at the field site in Grumeti, numerous cattle herds were seen near the river at the border and there was evidence of sustained trampling of the low-lying vegetation near to the river from where herds had been taken to water, which was not seen to the same degree at other sites. If these herds had been treated with insecticide which kills tsetse, this would explain why catches were low near to the river, but increased as the traps were further from the river. The significant relationship with elevation likely reflects distance from a river as the traps along the river were at the lowest points.

There were more significant relationships found with environmental variables within grids than within river transects. Catches at Ikoma were significantly related to Band 7, NDVI, LST and percentage tree cover. This is the site which is most shielded from external factors such as control as it was furthest into a protected area and therefore indicates that environmental effects are most strong in the absence of human interventions. Elevation across this area did not differ much and the river used in the shapefile for analysis was dry at the time of sampling so the lack of significance with these factors might be expected. Inner Ikorongo was also a relatively protected area, although had been subject to recent disturbance due to the construction of a bridge in the sampling area, and was significantly correlated with NDVI and percentage tree cover. Catches at the North Ikorongo grid were significantly correlated with elevation, percentage tree cover and distance from a river, which was similar to the results for the North Ikorongo river transect. Although this site bordered a farming area like Grumeti, the rivers in the North Ikorongo area did not travel along the border line as it does at Grumeti. The presence of insecticide treated cattle is therefore less likely to have the same impact.

There were slightly fewer significant relationships with *G. swynnertoni* catches from the four grids as there was for the river transects, although it was for different factors. For the grids with all grids combined the only significant factor was percentage tree cover, which was opposite to the trend seen for *G. pallidipes.* Within grids this was however not

significant. It appears as though the higher catches seen in Ikoma, which are likely to have occurred due to more wild hosts and lack of control, are biasing the relationship as without Ikoma included the relationship for all grids would no longer be significant.

2.4.3 Combined analysis

Combining the river transects and grid sampling data points should provide a greater power to examine relationships. For *G. pallidipes* this showed significant relationships for all the environmental variables apart from NDVI. This highlights the importance of environmental variability on *G. pallidipes* abundances. Although NDVI is often correlated with *G. pallidipes* catch, the levels of 'green-ness' could have varied across the two seasons sampled in depending on new growth or drought and this may have made the data points spread across a greater range of values when the two time-points are viewed together.

For *G. swynnertoni*, only LST and percentage tree cover were significantly related to average catch. This again reinforces that *G. swynnertoni* is capable of surviving over a greater range of conditions, however LST is still a critical factor. The relationship with percentage tree cover again highlights the difference to *G. pallidipes*.

2.4.4 Method for determining tree cover

Percentage tree cover was the environmental variable which has the highest number of significant relationships with individual grids. The manually calculated method provided greater adj. R^2 values when relationships were significant, however is much more time-consuming and could be vulnerable to the bias of the researcher knowing what the site looks like in real life. Using the Hansen file provides a much quicker way of extracting data, and was significantly correlated with the manual method which proves it's accuracy.

However, the Hansen tree cover file is from 2000 and is not regularly updated and would therefore not reflect any effects caused by burning of vegetation, intentional or wild. It is also significantly correlated with both Band 7 and NDVI, which can be collected more frequently from Landsat images so those are perhaps more useful.

2.4.5 Pteridine analysis

The grid sites studied were deliberately chosen to provide a range of different scenarios. Ikoma and Inner Ikorongo were inside protected areas, whereas Grumeti and North Ikorongo were on borders of protected areas. It was hypothesised that at Grumeti and North Ikorongo the tsetse populations may be subjected to a higher mortality due to the presence of humans and cattle and that may be seen through a different age structuring in the populations. It would be likely that an area subject to a high mortality would have fewer

young flies as those born there may not survive very long, and are likely to be populated by older flies which had travelled in from another area. The pteridine analysis did show a significant difference between Grumeti and Ikoma for *G. pallidipes*, and in terms of potential insecticide exposure these two sites are likely to be the most contrasting so this theory may fit. However the fact that there was no significant differences between the other populations could indicate that levels of dispersal of flies in and out of populations are similar.

Although the relationship between ovarian age and standardised fluorescence for *G. pallidipes* flies was significant, there is still a level of variation within each age category which means that for example a fly with ovarian age of 30 days could have the same standardised fluorescence as a fly <8 days old. The slope of the relationship for *G. swynnertoni* was steeper which would reduce the level of uncertainty seen for *G. pallidipes*, however it appears that flies <20 days old by ovarian age category have very points below the standard curve, which could explain why it appears that there are fewer young flies in the *G. swynnertoni* populations. This is still however the first time this method has been applied for *G. swynnertoni* and shows potential for further use. For both species there is also likely a trapping bias (Vale, 1974; Hargrove and Ackley, 2015) as only stationary traps were used to collect flies which will not give the entire picture of the population structure.

2.4.6 Fat content

It was hypothesised that flies living in more suitable habitat will be in a better nutritional state and have a greater fat content. The results of this study for *G. pallidipes* however appear to show the opposite. Flies caught in border areas of Grumeti and North Ikorongo were found to have a significantly greater fat content than those collected in the Serengeti National Park or deeper inside Ikorongo. This is a similar result to the result gained by Bursell, (1966) where although the habitat is less likely to be suitable in these areas, the flies appeared to have a greater fat content. It is possible that if control is ongoing in these border areas, it is biasing the type of flies caught in traps. Inside the Serengeti National Park or Ikorongo, there will not be an increased mortality from human impacts, however on border areas if cattle are treated with insecticide it is likely that very hungry flies may have attempted to feed on these cattle and subsequently died, meaning this section of the population is not being caught in traps and only gives a select picture of the population. Although age structure makes it appear as though flies are equally dispersed, age structure is likely showing a long-term effect and due to potentially high rates of mortality, fat content is showing a more local, short-term effect. As *G. swynnertoni* are likely to be

dispersing further if this theory is correct then it is likely that the effect will be seen less strongly, and while *G. swynnertoni* flies caught on the border did still have a higher fat content, it was not significantly different to flies caught in protected areas.

2.4.7 Limitations and solutions

The limitations of this study were mainly logistical. Satellite images were used to select areas within sites which had a range of vegetation cover, from highly vegetated riparian areas, to open savannah areas. The initial plan had been to set out grids with a square 4 by 4 trap set-up of equal spacing from each other, however due to a lack of road access, traps were mostly set out on foot and due to the presence of wild animals and/or difficult terrain, the shapes of the grids were modified to be closer to the roads. A comparable even spread across areas with different types of vegetation for each was therefore not always achieved.

A technical limitation of this chapter was that the pteridine method could only be calibrated using female flies, as their age could be independently determined by ovarian dissection. To be able to calibrate males from the area, pupae would have had to have been collected, hatched and kept alive until they reached set ages when their pteridine level could be determined, such as done by Lehane and Hargrove (1988) on Redcliff Island in Lake Kariba, Zimbabwe. Using only females may have introduced a bias into the age structure of the population, as female flies tend to live longer and travel further distances (Vale, Hursey, Hargove, *et al.*, 1984).

Percentage tree cover was a factor which was highly correlated with *G. pallidipes* abundance, however establishing this manually using QGIS is very time consuming and the Hansen images can be unreliable over time. One method which could provide a quicker solution for larger areas would be to use supervised classification techniques (Fornace *et al.*, 2018). This method would be first applied to identify vegetation suitable for tsetse over a small area to optimise the technique before the image classification is then applied over the entire area of interest. This method was successfully used by Carrasco-Escobar *et al.*, (2019) to detect mosquito larvae habitats.

2.5 Conclusion

Environmental variables generally correlated well with *G. pallidipes* catches over small scales (~100 m) inside protected areas. This is much less likely in border areas, and in some border areas where a high mortality is expected, the patterns appear to be the opposite to what would normally be expected for the species. As expected*,* catches of *G. swynnertoni* did not correlate with many of the most commonly examined environmental variables, apart from percentage tree cover, which re-confirmed that the species is more commonly found in and suited to open areas.

The age structure of tsetse populations mostly did not differ in each area for either species, perhaps indicating a similar level of immigration/emigration from each population, despite differences in habitat. Fat content for *G. pallidipes* across different habitats did show significant differences, however, this is likely due to sampling bias and/or differences in mortality.

Overall, these findings show that *G. pallidipes* abundance could be predicted well, when inside protected areas. It also indicates that it would be more difficult to predict in border areas, and there are likely external factors such as vector control which could be obscuring the impact of environmental variables on the distribution and abundance of tsetse.

Chapter 3: Use of insecticide-treated cattle to control tsetse at the interface of conservation and farming areas

3.1 Introduction

Treating cattle with insecticide is a method of vector control which has been utilised for several decades (Hargrove *et al.*, 2000). As part of a collaborating project, a questionnaire was conducted (Allan, 2019) with livestock keepers in the same areas of the Serengeti where tsetse sampling had been undertaken. The livestock keepers were asked about their knowledge and practices of vector-borne disease and vector control, and they reported a high level of use of insecticide on their cattle (Allan, 2019). This could explain why in areas where cattle are present, the predictive tsetse abundance model (Lord *et al.*, 2018) does not work well as a high mortality due to vector control is not incorporated into the model.

Malele *et al.*, (2011) states that high numbers of insecticide treated livestock have altered the ecology of tsetse and their hosts and that this has led to the disappearance of species of tsetse in Tanzania.

3.1.1 Impact of insecticide on tsetse densities and trypanosomiasis

One of the first reported uses of insecticide applied to cattle is from Thomson (1987). The author investigated the effect of deltamethrin on *G. pallidipes* and *G. m. morsitans* when applied to cattle as a spray or as an impregnated ear tag. The study took place at the Rekomitjie Research Station in Zimbabwe. One ox was sprayed with Decatix – a deltamethrin based dip made up to 0.0046%, and a second ox had two ear tags applied which were impregnated with 0.4 g deltamethrin. A third ox was left untreated to provide a control. Flies were caught during twelve sessions over eight weeks from the sprayed and untreated cattle, and during sixteen sessions over nine weeks from the tagged cow and the control cow. Any flies which landed on the oxen were caught with a hand net, before being placed into glass tubes and transported to the insectary where their condition was assessed on arrival, and 48 h later, using descriptions of knockdown from Quinlan and Gatehouse (1981). A bioassay was also performed on the tagged ox to determine which part of the cow gave the highest mortality to flies which were exposed for 30 s. Mortality of flies which had been in contact with the sprayed ox was 95% during the first two weeks after treatment, and a knock down within 4 h of exposure was at least 70% for eight weeks. For flies which had been in contact with the tagged ox, knock down was never higher than 41% and mortality never higher than 16%. The difference in results was attributed to the distribution of insecticide. The insecticide on the sprayed cow was assumed to be relatively

evenly distributed, however the bioassay on the tagged cow showed a higher mortality on the head, neck and the top of the front legs, indicating that this is where the insecticide concentration is higher. The number of cattle used in this study is however very low which means that results would need to be compared with other studies to be convincing.

A paper by Thompson *et al.*, (1991) examined results from five studies, including three larger scale studies. The first involved a herd of 2,403 cattle in Zimbabwe which were normally all dipped weekly for tick control with an acaricide called dioxathion, an organophosphate. Between January 1983 and March 1984, the dip in one out of the five tanks was changed to deltamethrin and 331 cattle were treated with this instead. Over the study period trypanosomiasis in the deltamethrin-treated cattle reduced from 8.9% to 3.9% while the acaricide-treated cattle cases increased from 1.7% to 3.5%. This suggested that the tsetse population around the deltamethrin cattle had decreased, but without recorded tsetse densities before and after the results cannot definitely be linked. A separate study in the Thompson *et al.*, (1991) paper from Zanzibar found 46% of cattle in an area of 20 km^2 were infected with *T. congolense* or *T. vivax.* There was an average catch of 1 tsetse per target per day, at 26 locations. The cattle were treated with a trypanocides but within two months, 38% of 150 cattle were infected again. A total of 700 cattle were then treated with a pour-on of 1% deltamethrin. Within 37 days the apparent tsetse densities had dropped to zero and were not caught during a two month observation period. None of initially sampled cattle had been re-infected during the same period. While this study was on a larger scale than that from Zimbabwe, Zanzibar is an island with little risk of reinvasion from adjacent uncontrolled areas. Thus the results from Zanzibar did not provide strong evidence that use of insecticide-treated cattle would be effective in larger mainland areas of Tanzania. The largest separate study in the paper reports on 26,244 cattle in Zimbabwe which were dipped fortnightly with deltamethrin at a concentration of 37.5 ppm. Four months after dipping began, another 11,667 cattle in surrounding areas began being treated with pour-on 1% deltamethrin. The number of trypanosomiasis cases per month in the dipped cattle dropped to zero from approximately 10 within three months from the start of dipping, and from approximately 10 to zero within six months for the pour-on treated cattle. These three larger scale studies have shown that trypanosomiasis cases have fallen, however it is only on Zanzibar where they examined actual tsetse densities.

Bovine trypanosomiasis was also reduced in Zambia with deltamethrin dipping (Luguru, Bennett and Chizyuka, 1993). Three herds were sampled for trypanosomes, treated with a trypanocide and then dipped in a 0.00375% deltamethrin wash. Two herds were dipped

every two weeks and the third was dipped weekly. Before dipping began, rates of trypanosome positive cattle were between 27 – 51%. During dipping the rates declined to 1 – 3%.

At the site which is now Saadani National Park in Tanzania, ITC was frequently used when the Mkwaja ranch was in operation (Fox *et al.*, 1993). The ranch opened in 1954 and trypanosomiasis was initially controlled well with prophylactic trypanocides. However, by 1988 the trypanosomes appeared to be developing resistance. Aerial spraying, ground spraying and SIT had been attempted previously. Aerial spraying was found to be too expensive, ground spraying too environmentally damaging, and the SIT trial could not maintain results with a 1 km vegetation-cleared barrier as re-invasion occurred (Williamson *et al.*, 1983). In 1989, dipping of almost eight thousand cattle with 0.00375% deltamethrin began, initially every two weeks but reduced to weekly intervals during rainy periods. Apparent densities of tsetse in June of 1989 were recorded as 142 but fell quickly after dipping began. *Glossina m. morsitans*, *G. pallidipes* and *G. brevipalpis* were reduced by almost 100%, 90% and 70% respectively by April 1990 and remained roughly at these levels until March 1991 (Fox *et al.*, 1993). This paper is one of the best large-scale studies of ITC which include data on changes in the population density of tsetse following the treatment of cattle.

Deltamethrin was also tested at a ranch in Uganda which had previously only been treating for ticks using dioxathion (Okello-Onen *et al.*, 1994). Between April and July 1989, a herd of 700 cattle was treated with 0.00376% deltamethrin every two weeks at three time points in the season when there was light rain, followed by weekly dipping during the heavier rainy season. A control herd of 200 was treated only with dioxathion. Before deltamethrin treatment began, the average *G. pallidipes* catches per trap per day at the treatment ranch and the control (non-intervention) ranch were 6.50 and 6.0 respectively. At the treatment camp the average daily catch per trap dropped to 0.20 during the fourth week after the second treatment, and during the following eight weeks no tsetse were caught. At the control ranch the number of tsetse fluctuated and towards the end of the trial did drop to 0.2 tsetse per trap per day, however the overall average across the treatment period was 4.83, which is significantly higher than the overall average of 0.81 for the treatment ranch. The drop to 0.2 at the control ranch was attributed to a very dry period of weather. This paper also shows a frequent use of deltamethrin on a large number of cattle having a significant impact on tsetse densities.

Several studies in the 1990s reported successes of controlling tsetse numbers using insecticide treated cattle. Leak (1995) describes how in the Ghibe valley, Ethiopia, between 1991 and 1993, 2000 - 4000 cattle were given a monthly treatment of the pyrethroid cypermethrin applied as a pour-on. This resulted in a 93% reduction of the apparent density of *G. pallidipes,* and a reduction of 83% of apparent density of *G. m. submorsitans*. This was associated with a decrease in the prevalence of trypanosomes detected in the cattle of over 74%. A study in Burkina Faso (Bauer *et al.*, 1995) measured tsetse density changes after 1500 to 2000 cattle were treated with deltamethrin spot on. The cattle were initially treated at monthly intervals for four months, before decreasing treatments to every two months. Initial tsetse densities were 54.2 flies per trap per day, consisting mostly of *G. m. submorsitans,* and after 11 months of treatment the densities had been reduced to 0.06-2.0 flies per trap per day, which mostly consisted of *G. palpalis gambiensis.*

In Eastern Zambia a trial of cyfluthrin was undertaken using a pour-on to try to treat just less than 100% of the total cattle population for one year (Van Den Bossche *et al.*, 2004). Treatments were given every seven weeks but the number of cattle treated varied. The average for the trial period was 61.6% out of all the cattle in the area. The maximum average was 88.3% ±10.2% and the lowest average was 33.6 ± 1.1%. Prevalence dropped from approximately 10% one month after treatment began, to 0 cases after nine months. In this area, sales of trypanocidal drugs declined from 13,134 sales in the year preceding treatment to 3,738 in the year when treatment was taking place, indicating farmers confidence in the insecticide treatments to protect their cattle. Tsetse densities were not reported in this, however other studies had started to build evidence that the reducing tsetse densities and reduced cases of trypanosomiasis were linked.

In Ethiopia, the effect of treating cattle with a pour-on formulation of deltamethrin was compared against deltamethrin-targets targets against *G. pallidipes* (Bekele *et al.*, 2010). Two sites were set up in the southern rift valley of Ethiopia. At one site, 460 targets deployed at a density of 4 per km^2 and at a second site, 409 cattle were treated with 1% deltamethrin pour-on. Both were effective with the catch of tsetse from monitoring traps declining by 88.9% and 94.9% respectively.

Gouteux *et al.*, (1996) describes the treatment of cattle with pour-on formulations of flumethrin or deltamethrin to control *G. fuscipes fuscipes* in the Central African Republic. Four herds, each of about 40 cattle, were treated every three weeks with flumethrin during the rainy season and with deltamethrin every six weeks in the dry season. The use of

insecticide on these herds appeared to cause no changes to the apparent density of tsetse, or have any impact on the prevalence of trypanosomes in the cattle. The authors concluded that in this setting of small herds which are geographically isolated such as is the case of the Fulani farmers, the use of insecticide treatments on cattle is not sufficient to control tsetse and trypanosomiasis. Although this paper examines the impact on a riverine species of tsetse, it is important to examine cases where ITC had failed to see if there was any transferable lessons to be learn.

Another study where ITC was unsuccessful for its desired purpose was in Zimbabwe (Warnes *et al.*, 1999). A barrier of insecticide treated targets had been in operation with 4 to 5 targets per km^2, alongside the use of 5400 cattle treated with a deltamethrin dip and pour on every two weeks and every month, respectively. The barrier of targets and cattle was approximately 20 km wide from the reinvasion front. Tsetse catches and trypanosomiasis incidence was monitored for roughly 8 months, then the treated targets were removed leaving ITC as the sole vector control method. Within the first month after removal of the targets the numbers of tsetse and the distance tsetse were caught from the reinvasion front increased, and gradually continued to rise over the following seven months. Trypanosomiasis prevalence had also increased so the targets were re-deployed, causing the tsetse catches and trypanosomiasis prevalence to decrease. The authors stated that their study was designed to see if ITC could prevent tsetse invasion, rather than control populations to manageable levels. They concluded that ITC was not capable of stopping tsetse invasion due to uneven distribution of cattle along the invasion front.

Overall, there is substantial evidence that the use of ITC can suppress tsetse populations, as long as the insecticide treatment is used in high enough quantities. This is likely to be especially effective on border areas such as in our study areas around the Serengeti, as reinvasion would not be occurring from all sides, and is therefore more easily maintained.

3.1.2 Knock down and mortality caused by insecticide to tsetse

To confirm that insecticide is reducing cases of trypanosomiasis and densities of tsetse by causing fly mortality rather than through repellent effects, many bioassays and observational studies have been performed. In the Thompson *et al.*, (1991) paper, the author examined knock down of *G. pallidipes* exposed to one cow treated with a deltamethrin spray (46 ppm) and one treated with a deltamethrin pour on (1%). Wild flies which landed on the cows at periods over the following ten weeks were collected with hand nets and taken to an insectary to monitor. Over the first two weeks of the study, >90% of flies were knocked down by both types of treatment. At six weeks post treatment,

knock down of flies for the pour on was 40% and 49% for the spray and these had declined respectively to 16-20% and 0-7% by week 10.

Okiria and Kalunda (1994) also examined knockdown with three species of tsetse. One cow was dipped in 0.00375% deltamethrin and another cow was dipped in water. The susceptibility of wild *G. fuscipes fuscipes* and colony-reared *G. morsitans* and *G. pallidipes* was examined. The flies were exposed to the cow and allowed to feed. Knock down was observed at 30 min post blood meal, and mortality measured at 6, 24 and 48 h after feeding. Tsetse were exposed to the cow on days 0, 3, 7, 14, 17, 21, 24, 28 and 31 days after treatment. All flies which were exposed to the treated cow within 21 days were knocked down within 30 min. At 31 days, *G. f. fuscipes* did not show any knock-down, however *G. pallidipes* and *G. morsitans* had 80% and 40% knocked down within 30 min respectively. The main species caught in the Serengeti and Saadani sites in Tanzania was *G. pallidipes* so this study gives confidence that *G. pallidipes* is significantly affected by pyrethroids, and for deltamethrin this high effect lasted for a month.

Knock down of wild caught *G. pallidipes* and *G. m. morsitans* was investigated against three different types of insecticide in Zimbabwe (Vale, Mutika and Lovemore, 1999). Deltamathrin, alphacypermethrin and cyfluthrin provided at least 50% knock down for 5-24 days in hot months, and 24-55 days during cooler months. Subsequent modelling found that 4-21 annual applications of insecticide would provide a good level of tsetse control in areas greater than 1000 km^2 for at least 10 km from the area where the tsetse would be re-invading from, assuming that the insecticide-treated cattle made up at least 50% of the tsetse diet.

Tsetse are attracted to cattle making them an ideal bait for tsetse (Vale, 1974), but evidence that the presence of insecticide on cattle would not make them less attractive was needed. It was also important to confirm that the reason for reduced trypanosomiasis in tsetse controlled areas was due to the death of tsetse, or if not, perhaps the insecticide was affecting the ability of tsetse to probe and feed. If it was the latter, a gap in treatment of insecticide could cause a sharp rise of cases if tsetse are not dying, and are then given a chance to feed again. An experiment in Kenya (Baylis, Mbwabi and Stevenson, 1994) treated groups of two to four cattle with either Spot On – a deltamethrin-based insecticide or Ectopor – a cypermethrin based insecticide, and one control group which was untreated. One animal per day would be placed within an incomplete circle of electric nets which sampled a proportion of the tsetse that visited and fed on the cattle. Their experiment

found that a pour-on of deltamethrin or cypermethrin did not significantly affect the number of *G. pallidipes* attracted to cattle, or the proportion of flies which had fed from it. The authors acknowledged there was however large variation within their data, and recommended that if this variation was to be overcome, a greater number of host animals and experimental days would be needed. Torr *et al.*, (2007) also found no evidence that deltamethrin was repellent to *G. pallidipes* or *G. m. morsitans*.

3.1.3 Impact of ITC use on cattle health/productivity

If the use of ITC would help improve the health of farmers herd, this would make its uptake more likely. There are several papers which report on the impact of ITC on cattle health and productivity. For instance, Luguru *et al.*, (1993) stated that cattle owners had commented that there was a higher number of calves born and greater milk production during dipping which had not been seen before dipping began.

At Mkwaja ranch, the calving rate rose from 58% to 77% following the introduction of dipping with deltamethrin. Weaning weight also increased from 124 kg to 145 kg in the same time period. Using a three year average from before deltamethrin treatment began, mortality attributed to trypanosomiasis dropped from 0.4% to 0.1% during treatment, abortions dropped from 0.75% to 0.33%, and pre-weaning mortality dropped from 14.4% to 4.6% (Fox *et al.*, 1993).

In Burkina Faso, the study by Bauer *et al.*, (1995) reported an improvement in cattle measurements as tsetse densities decreased following treatment with insecticides. They recorded a weight increase of calves aged 6-12 months from 122.3 kg to 213.6 kg over a year from October 1993 to October 1994. An average daily weight gain of more than 400 g was measured over four months from April 1994. In a second study from Burkina Faso, (Bauer *et al.*, 1999) reported that incidence of AAT declined from 30% before insecticide treatment to 5%. This was maintained during the rest of the intervention period up to 1996. The percentage of abortions also dropped from 66% in 1993/94, to 10.1% in 1996/97, and milk offtake increased fivefold during rainy seasons and elevenfold in the dry season over the same time period. In this study, 1500 insecticide impregnated targets were also utilised during the dry seasons alongside insecticide on the cattle and helped reduce the tsetse populations by 90%.

As part of the tsetse control programme in the Ghibe valley in Ethiopia (Rowlands *et al.*, 1999), calf mortality was reduced by 57% and the body weight of adult males increased by

8% following mass treatment of cattle with pyrethroids. The body weight of cows and the calving rate however did not increase, despite a reduction of tsetse densities of 63%.

3.1.4 Restricted application of ITC

Although ITC is a relatively cheap method of vector control, the cost may still be prohibitive for poor cattle owners. Studies in Zimbabwe have found that tsetse do not land equally on all parts of cattle to feed, but instead feed more on the belly and lower legs, and more often on older cattle (Torr and Hargrove, 1998). It was suggested that ITC could be more cost-effective if the whole animal did not need to be treated, and farmers could therefore treat more animals with the same amount of insecticide (Torr, Maudlin and G. A. Vale, 2007).

Torr *et al.*, (2007) exposed wild, female *G. pallidipes* to cattle which had been treated with either a 1% pour-on or a 0.005% suspension concentrate of deltamethrin. The insecticide was applied to the whole body, belly and legs, legs, front legs, middle and lower front legs, or lower front legs. Knock-down achieved when the whole body was treated was highest, however the knock-down percentages achieved when only the legs and belly had been treated were still relatively similar. The other treatment options did not compare as favourably. The authors calculated that treating only the legs and belly at 2 weekly intervals, rather than treating the whole animal monthly, would make cost savings of 40% and improve efficacy by 27%.

Restricted application of deltamethrin can be done on an individual animal by not treating the whole body, but application can also be restricted in the sense that only a percentage of the herd is treated. This was tested in south-eastern Uganda over an 18 month period (Muhanguzi *et al.*, 2014). Cattle were selected from 20 villages and a protocol of no treatment, 25%, 50% or 75% monthly treatment was applied to different herds. The paper does not report the impact of the interventions on density of tsetse but they did find that all levels of treatment (25-75%) provided similar levels of protection against AAT. The cost reductions of this method were also calculated and made it an attractive option for vector control, as treating 25% of the cattle would cost a village US $1.72 per animal per year, compared with $5.17 per animal per year to treat 75% of the cattle (Muhanguzi *et al.*, 2015).

One concern about the use of ITC is the impact that insecticide may be having on non-target species. Vale *et al.*, (1999) found concentrations of up to 0.15 ppm alphacypermethrin in dung from the treated cattle for almost two weeks after the cattle

had been treated. Approximately 500 dead or dying dung beetles were collected from inside or very near to the dung of a cow the day after it had been treated with alphacypermethrin. Vale *et al.*, (2015) found that restricted application of insecticide can mitigate these risks. Their results found that dung was likely contaminated with insecticide by a combination of cattle ingesting the insecticide as they licked themselves and the dung coming into contact with the body when the cattle defecate. By applying insecticide to only the legs and belly, the numbers of dead dung beetles found on dung pats was reduced to negligible levels.

3.1.5 Tick resistance

Although insecticides may be useful for reducing tsetse numbers and other disease vectors, when used inappropriately they are leading to resistance in ticks, and possibly worsening the rates of tick-borne disease. Important tick-borne diseases include anaplasmosis, babesiosis and East Coast fever. Tick resistance to insecticide has developed largely because the doses needed to kill ticks are greater than those needed to kill tsetse. In the 1970s and 1980s, tick control was very successful in South Africa and Zimbabwe. Here, large areas of each country were treated with sprays by government teams. Unfortunately, various political, social and economic issues focused the governments priorities elsewhere and capacity for control programmes was vastly diminished. Responsibility for control then shifted to the livestock owners themselves (Eisler *et al.*, 2003). Since then, many studies have reported increasing acaricide resistance.

A study of *Rhipicephalus microplus* in South Africa found one of three populations to be highly resistant to synthetic pyrethroids and moderately resistant to pyriprol (Lovis *et al.*, 2013). A recent study in Zimbabwe reported 38% of *Rhipicephalus microplus* ticks collected from cattle at a communal dipping tank had complete resistance genotypes against amitraz (Sungirai *et al.*, 2018).

A study of *Rhipicephalus appendiculatus* and *Rhipacephalus decoloratus* in Uganda found many acaricide resistant ticks. Resistance to synthetic pyrethroids was found in 90% of tick populations. Many ticks were also resistant to multiple acaricides. This includes 60% and 63% of the above mentioned species being resistant to two times the discriminating dose of cypermethrin and deltamethrin. Of the farms studied, 40% of them had rotated their acaricide but used the same molecule in a different brand, and 7% of farms had mixed two different formulations together (Vudriko *et al.*, 2016). In the absence of proper education on how to rotate products effectively, farmers have attempted to come up with solutions which may have exacerbated the problem.

3.1.6 Integrated control

The areas where *Trypanosoma* spp. are found substantially overlaps with areas where tick-borne pathogens are found (Figure 30). This can have significant impacts on mortality; for cattle infected with ECF, the likelihood of death was six times higher if co-infected with *Trypanosoma* sp. (Thumbi *et al.*, 2014). Despite the risk of acaricide resistance developing in ticks, if managed properly there is great potential to integrate the control of both vectors, and would be especially useful in resource-poor settings.

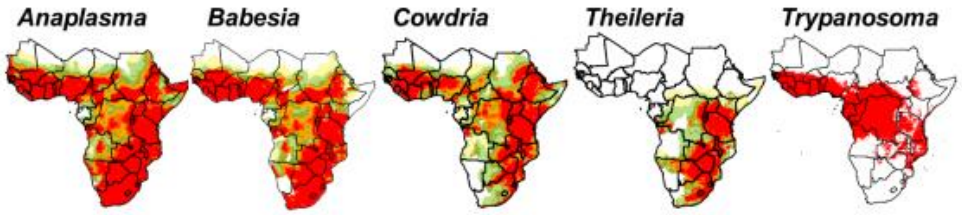

Figure 30. The distribution of cattle diseases pathogens spread by ticks and tsetse (Minjauw and Mcleod, 2003).

If young cattle are over-treated with insecticide for tsetse or tick control, they will not be able to build up a natural immunity to tick-borne disease (Eisler *et al.*, 2003). If for some reason there was a gap in application of insecticide, whole herds would be susceptible to disease and mortality rates likely to be high if they became infected. Integrated control of tsetse and ticks should aim to maintain endemic stability where infection can occur without a high rate of clinical disease (Eisler *et al.*, 2003). There are several ways this can be done. As tsetse are more sensitive to insecticides than ticks, it should be ensured that the treatment times are as long as possible without leaving the animals exposed. To allow young animals to gain immunity, only the older cattle in a herd should be treated, and as tsetse preferentially feed on older cattle (Torr, Maudlin and G. A. Vale, 2007) this should not impact on tsetse control. Restricted application of insecticide is also useful for reducing the impact on ticks. For example, *Rhipicephalus appendiculatus* is associated with attaching to the ears of cattle (Kimaro, Toribio and Mor, 2017), whereas tsetse feed mainly on the legs so insecticide application should be focused on the belly and legs (Torr, Maudlin and G. A. Vale, 2007).

3.1.7 Modelling of insecticide treated cattle as a method of vector control

The use of insecticide treated cattle has been shown to be effective in most scenarios, but occasionally it has not been sufficient as a standalone tool to reduce tsetse densities and/or trypanosomiasis transmission. To allow end-users to determine if ITC would be effective in a specific area, and to calculate under what conditions it would need to be applied, models have been developed to assist with designing control strategies.

One of the first was developed by Vale and Torr (2005) and was used to calculate the cost and effects of SIT and ITC. The parameters they used for SIT were weekly releases of three sterile males released for every one wild male and for ITC this was 3.5 cattle per km^2. For a population of 2500 males and 5000 female tsetse per km^2, ITC control would take 187 days to eliminate the population and SIT would take 609 days. The authors state that even under the most favourable conditions, SIT would still cost 20 times more than ITC. ITC is therefore not only a much cheaper option to implement, it would also be able to eliminate populations in a much faster time, and reduce the risk of trypanosomiasis so less trypanocidal drugs would need to be bought which would also save the cattle owners money.

Torr and Vale (2011) then investigated the impact that distribution of cattle in an area may have on the effectiveness of ITC. They found that an uneven distribution of cattle over a small scale was not likely to alter the efficacy of ITC, however any areas over 3 km^2 wide without any cattle in should have insecticide treated targets placed within them to maintain impact. Most of the areas around the Serengeti and the bordering game reserves have cattle dispersed every few km.

Hargrove *et al.*, (2012) also used models to show that in areas of Uganda where wild hosts are scarce, Rhodesian HAT could be more effectively controlled with ITC, rather than trypanocides.

Cost is one of the most important factors when considering different tsetse control measures (Shaw *et al.*, 2013). To compare the costs of different methods, Shaw *et al.*, (2013) modelled the cost of controlling trypanosomiasis over a hypothetical area of 10,000 km^2 in Uganda. The cheapest method of control was ITC followed by traps, SAT, then SIT. Continuous control was recommended as total elimination in some areas can be very difficult to achieve, especially when the risk of reinvasion is high. Kajunguri *et al.*, (2014) used a model of the transmission of *T. b. rhodesiense* in tsetse to support results which suggest that ITC is capable of reducing tsetse density and *T. b. rhodesiense* incidence in humans, and that ITC is best applied in a restricted manner to improve cost-effectiveness.

3.1.8 Public perception of ITC

Public understanding and acceptance of control methods is vital for successful operations. One of the first studies to assess farmers perceptions of the use of ITC was conducted in southwest Ethiopia (Swallow, Mulatu and Leak, 1995). Survey results were gathered from 166 participants, from the same study area and time reported on by Leak (1995). For

almost two years, the pour on treatments were provided free of charge, then after December 1992 a charge was introduced for each animal. When the pour-on was free, 99% of participants had used it. After the charge had been introduced, this dropped to 67% of participants. This drop was attributed to the cost but also to the fact that the number of tsetse had declined and farmers perceiving that there was a lower risk of trypanosomiasis. Geographical distance to a crush (structure used for restraining cattle) was found to be significant, as was the number of oxen and cows in the herd. If a house was double the distance from the crush as the mean household, they were 36% less likely to travel to the crush, and if a household had 10% more oxen, or 10% more cows than the average household, they were 2.8% and 4.4% respectively more likely to get the treatment. The reduction in demand for the pour-on following the cost increase did not affect the impact on tsetse or incidence of AAT, suggesting that even though participation declined, a sufficient number of farmers were still using the insecticide and thereby controlling the tsetse population.

A questionnaire was used during the tsetse control programme in Yale, Burkina Faso to assess the impact on cattle health attributed to tsetse and trypanosomiasis control, which could not be gathered by health monitoring (Kamuanga *et al.*, 2001). Farmers perceived that generally a reduction of trypanosomiasis cases led to the amount of milk each cow was producing to increase. Women are normally responsible for milking, and the gross income of the women increased from US $1.60 to US $3.25. Over 90% of participants believed this increase was due to improved animal health resulting from trypanosomiasis control. Almost 90% of participants were also aware that fewer cattle were dying from trypanosomiasis. Farmers reported that their offtake (proportion of animals sold in a year) was higher before tsetse control began, and as control was underway there was an increase in the numbers of 2-3 year old female cows as it was possible for farmers to build up their herd sizes. Farmers did however continue to use preventative trypanocidal drugs even when they were aware that the risk of trypanosomiasis had become much lower.

An investigation into the knowledge and attitudes of farmers in Uganda about tsetse and trypanosomiasis control was undertaken by (Magona, Walubengo and Olaho-Mukani, 2004). A questionnaire was done in two districts, one of which was Busia district, where Okiria *et al.*, (2002) described a vector control programme involving the treatment of ~1000 cattle with insecticide. Awareness of trypanosomiasis as a problem was high at 87.6%. Despite this, only 18.5% of participants across the two districts were aware of pour-on applications as a method of control, in comparison to 76.5% who were aware of trapping

tsetse. Participants were then given a selection of options of what type of tsetse and trypanosomiasis control they would prefer to participate with in the future, and only 25.9% said that they would be willing to pay for and use pour-ons, in comparison to 69.1% who said they would be willing to pay for prophylactic and curative trypanocidal drugs. The most highly preferred option would be for the government to provide traps and for the cattle owners to maintain and deploy them themselves. This preference was believed to be based on the fact that the government had previously deployed traps in southeastern Uganda as a means of controlling HAT, and the community felt the government had the resources to continue to provide it.

Again in Uganda, interviews were undertaken to determine insecticide usage and reasons farmers may or may not be using it (Bardosh, Waiswa and Welburn, 2013). In one area where HAT is considered high-risk, insecticides being bought which are effective against tsetse and ticks were bought in much lower volumes than those effective against ticks only. Factors affecting insecticide choice included the farmers knowledge of the disease, brand recognition, price and availability. Stakeholders considered restricting the market for products only effective against ticks to help improve tsetse control.

3.1.9 Current ITC situation in Tanzania

A questionnaire was administered to 70 households in the Serengeti District near to the national park by Mwaseba and Kigoda (2017) to investigate knowledge, attitudes and practices about tsetse control in their area. Their results found that generally, farmers were not able to correctly identify tsetse flies, however they were mostly knowledgeable about tsetse control measures. Most (80%) respondents said that recommended control methods were easy to use, and 70% said the methods were effective, however 73% said that control methods were expensive and 70% said they had not actually used any method of control against tsetse. Of those who did carry out disease control, 58% were either spraying or dipping and 42% were using treatments (paper does not specify what is meant by treatment). This was in contrast to data collected by colleagues from the Roslin Institute who also carried out a questionnaire in the Serengeti District, where the majority of farmers said they did use vector control methods (Allan, 2019).

At a national stakeholders meeting held in 2017 in Arusha, Tanzania, the head of the tsetse control division (Mrs Joyce Daffa) reported that the government had been offering guidance and subsidies towards the purchase of insecticides in the country. There is little data on sales of insecticide in the country, although Stanley (2016) reported on revenue gained from a dipping fee of 3,680,000 TSH between 2012 and 2014 in Makundusi village,

Serengeti. As of February 2017, there were at least 15 functional community dips in operation across the Serengeti District (Appendix A.5).

At a further meeting with stakeholders in 2019 (Oliva Manangwa, pers. comm.), more information was gained from the Tanzanian government on recent developments. In the last two decades, the Government of Tanzania has been introducing policies regarding the methods and types of insecticide used to control tsetse and ticks. In areas where both ticks and tsetse are present, the government has advised pyrethroids should be used to treat their livestock. In areas where only ticks are present, non-pyrethroid insecticides such as Amitraz should be used. This policy was designed to reduce the likelihood of resistance developing in ticks.

It is compulsory by law that cattle should be treated with insecticide. The policy differs by season with animals being treated every 14 days in the dry season, and every 7 days in the wet season. However, it is not compulsory for the whole herd to be treated at these intervals and in practice farmers treat only a portion of their herds at each application time on the understanding that the treated portion will provide a herd protection.

Between 2006 and 2014 the government also had a policy of subsidising insecticides. The subsidy was applied through government agrovet shops and at the point of dipping. The value of the subsidy was 40% of the cost, with the livestock owner paying the remaining 60%. In real terms this equated to a cost of 200 Tanzanian shillings per animal to be dipped in a dip-tank. The Sokoine University of Agriculture (SUA) was commissioned to evaluate this policy in relation to tick-borne disease prevalence (Mbassa *et al.*, 2017). In their report, they find that the percentage of cattle that were dipped did not increase between 2006/2007 (8.8%) and 2013/2014 (3.23%), despite an increase to 47.2% in 2008/2009. Regarding knowledge of the subsidy, they found even in villages with community dips only 42.2% of study participants were aware of the subsidy. In villages without dips, awareness was even lower at 26.9%. They examined reported cases of East-Coast Fever, anaplasmosis, babesiosis and ehrlichiosis and did not find a reduction in tick-borne disease during the subsidy period however the data was mostly incomplete, which the author attribute to a lack of diagnostic or insecticide testing facilities and poor communication links between the ministry and individual districts. Despite the lack of evidence that the policy has been successful in reducing tick-borne disease disease, the government re-started its subsidy policy in 2018. There is no mention of tsetse or trypanosome prevalence in the report, and it is possible that this is where the policy has had a large but unrecognised impact.

The maintenance of the community dips where farmers could access subsidised control is the responsibility of both the livestock keepers and the government. In a document provided from the Tsetse Control Division, they report that many of the community dips are actually out of order, so it is possible that the balance and processes for carrying out maintenance are not working as effectively as possible. The government did however report that in the last few months a concerted effort has been made to repair these dip-tanks.

Apart from government spending, it would have been difficult to get details on insecticide sales and distribution as each district had a different supplier. However, there is now a national supplier called Farmbase, which could make gaining detailed information much easier for any research over the coming years.

3.1.10 Methods of detecting insecticide

The presence and effect of insecticide control can be monitored by assessing its impact on tsetse and cattle health and productivity. However, being able to quantify insecticide concentrations in intervention areas would give additional evidence to directly support monitoring activities. There are a variety of physico-chemical methods available to do this.

3.1.10.1 Gas Chromatography-Mass Spectrometry (GCMS)

One method of detecting and quantifying insecticide in samples is to use gas chromatography – mass spectrometry (GCMS). Molecules in a sample are separated in a gas chromatograph by their relative affinity for the stationary phase of the column depending on their chemical properties. As the sample comes out of the column, the different molecules will come out at different times which are then passed into the mass spectrometer. The mass spectrometer then ionizes the compound and the resultant fragments can be detected and used to infer the identify of the original compound. This method costs in the region of £50 per sample to analyse if sent to an external provider of analysis and/or requires a well-resourced and equipped chemistry laboratory. It is not appropriate for resource poor settings.

3.1.10.2 High Performance Liquid Chromatography (HPLC)

HPLC involves the sample being passed through a column in a liquid phase, where molecules then bind to the stationary phase in the column based on their polarity. The molecules which have the greatest interaction with the stationary phase will leave the column the slowest. The retention time that is seen with a peak for a known standard of insecticide can then be compared with unknown samples to confirm presence with the same retention time and quantify the amount present. HPLC is potentially more reliable

than GCMS, as HPLC does not heat the samples which occurs in GCMS, which could degrade the samples. HPLC is also an expensive option and costs in the region of £50 per sample if sent to an external provider, or ~ £3 if done in-house (Ismail *et al.*, 2016), not including high initial equipment costs. As with GCMS it is not appropriate for routine use in resource poor settings.

3.1.10.3 Insecticide quantification Kit (IQK)

As a way to monitor indoor residual spraying programmes for mosquito control in the field, an insecticide quantification kit was developed (Russell *et al.*, 2014). This is a colorimetric test with a solution which turns red in the presence of cyanopyrethroids. The concentration of pyrethroid is indicated by the intensity of the colour and the colour is a function of the concentration and the time that the reaction proceeds. This test involves three reagents, and involves the detection of cyanide ions from the insecticide released by alkaline hydrolysis. This costs approximately £1 per sample and can be performed by anyone with very little training required in resource poor settings (Russell *et al.*, 2014). This relatively simple method has the potential for routine use in tsetse control programmes.

3.1.11 Chapter objectives

Lord *et al.*, (2018) showed that the density of tsetse in farming areas bordering the Serengeti National Park were lower than expected. Results presented in this thesis (Chapter 2) also suggest that the density of tsetse is lower outside the conservation area despite suitable habitat being present. One potential explanation for the unexpectedly low numbers of tsetse is that livestock keepers are treating their cattle with pyrethroids. The study presented in this chapter had the overarching aim of quantifying the use of insecticides by livestock keepers using a questionnaire survey and analysing cattle hair for the presence of pyrethroids.

Although an estimation of insecticide use can be gained from a questionnaire there are several factors which may be providing an inaccurate picture of the true situation. For instance, even if insecticide use is reportedly high, farmers will buy the insecticide as a concentrate and this will require dilution if not using a community dip. This means correct concentrations may not always be applied, leaving their animals vulnerable. There is also the possibility that farmers will provide answers which they think will be well received, rather than what they are practicing. Lastly, recall of the precise timing and concentration of insecticide applied may be poor. To gain a better and more objective picture of the control situation, molecular methods will be used to quantify insecticide concentrations present in herds alongside asking them details about the last method of treatment that they used.

A questionnaire conducted by colleagues from the Roslin Institute (Allan, 2019) was biased as the group of farmers interviewed were involved in a longitudinal study about foot and mouth disease (FMD). They were originally recruited to that study because they had alerted livestock officers about the potential of FMD being in their cattle. These farmers were therefore demonstrably more aware of diseases affecting their livestock and as owners of large herds, wealthier. Accordingly, a cross sectional survey of randomly-selected herds in Serengeti district was conducted to gather information on the treatment of cattle with insecticides. By implementing a random method of selecting herds, the survey would include the resource-poor livestock keepers with smaller herds, as well as the larger, and richer farms. Together these would provide a better representation of the overall insecticide use in the area.

During this study, a short questionnaire was performed and hair was collected from cattle. The questionnaire asked about insecticide use on each individual animal as hair was collected. The hair was returned to LSTM where it was analysed for insecticide using GCMS,

HPLC and a colorimetric method (IQK) which had previously been developed for analysing pyrethroids applied to bednets or walls to control malaria mosquitoes (Russell *et al.*, 2014). Quantitative analysis of the amounts of insecticide on cattle hair was also complemented by bioassays to assess whether the quantities detected are likely to have an impact on tsetse survival.

Therefore, the objectives of this chapter are to:

- Gather information from a questionnaire survey of livestock keepers about individual animal practices within their herds
- Quantify insecticide use in the Serengeti study area
- Determine if this insecticide use is sufficient to impact on tsetse populations
- Develop a cost-effective method for measuring insecticide concentration on cattle hair

3.2 Methods

3.2.1 Methods for Serengeti sampling

3.2.1.1 Selection process

Data on village locations was collected for collaborating projects (Allan, 2019). Sample size was determined with the aim of determining trypanosome prevalence in herds near to the National Park and game reserves. Sampling was stratified to include only the villages whose centres were less than 5 km from the boundary of the Grumeti game reserve, Ikorongo game reserve and Serengeti National Park. Within this sample, eight villages were randomly selected. Village leaders were then visited and a list of all of the sub-villages within that village requested. Sub villages were further stratified as close or far from boundary. Within those which were close to the boundary, one sub village was randomly selected, and within those which were further away from the boundary, another was randomly selected another.

From sub village leaders a list of the households within each sub village was requested, and households which owned cattle were highlighted. From this highlighted list, three households to have their cattle sampled and to complete a questionnaire were selected, three households for questionnaire only, and three further households to provide reserve herds should there be any problems locating, questioning or sampling from the selected herds on the day of sampling. The final list of households covered the majority of the area of interest at a range of distances from the protected border (Figure 31). The list of selected households was returned to the sub village leader so that they could contact the head of household and inform them of the research and ask if they were happy to participate in the study. Consent forms were provided on the day of sampling.

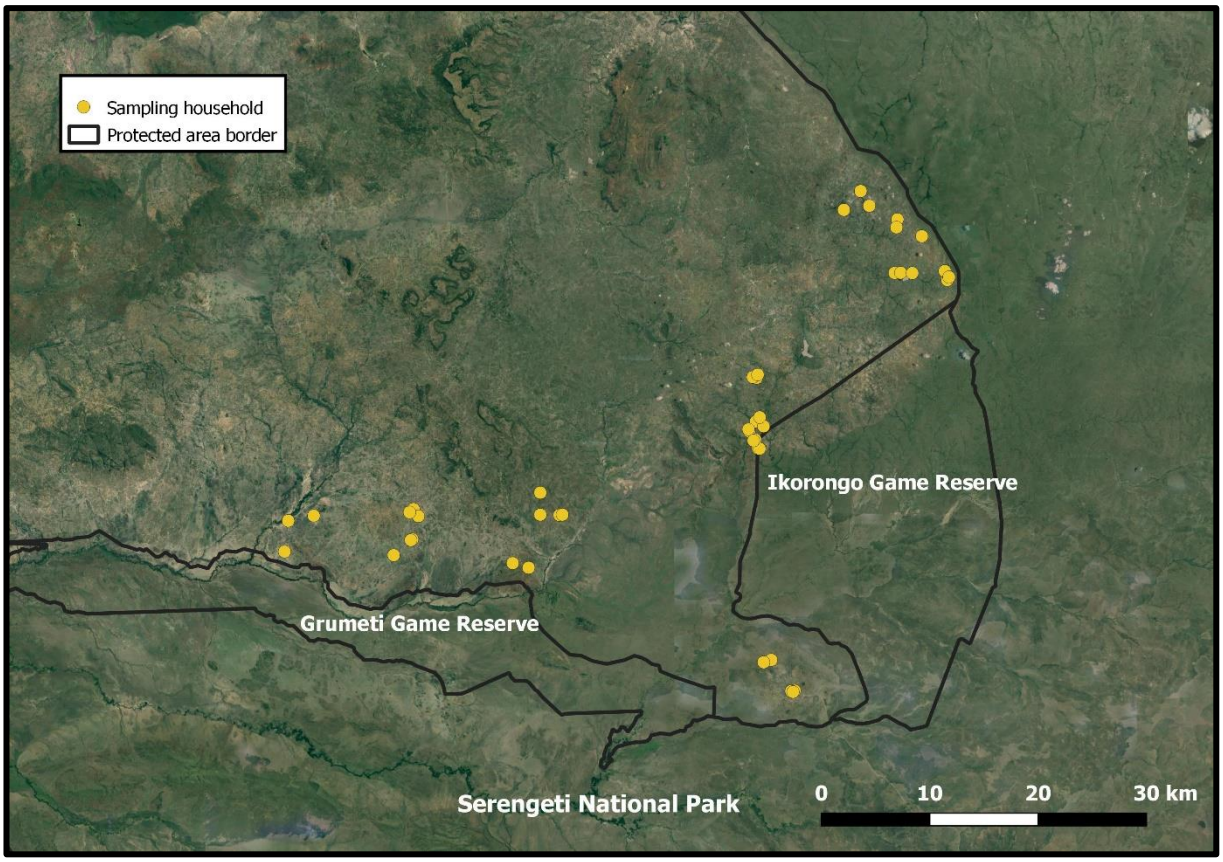

Figure 31. Map of cattle sampling households

3.2.1.2 Sampling

Sampling began by 06:00 hours to ensure cattle had not already left to graze. As part of the ZELS project a variety of samples were taken. Blood samples were taken from the ear capillary onto FTA cards and another from the jugular into both whole blood tubes and PAXGENE tubes. Tick samples were collected if they were of the species *Rhipicephalus appendiculatus* which are found on the ears.

The aim was to sample cattle at three households per day. The maximum number of cattle sampled per herd was 20; if herds were smaller than this all cattle would be sampled. If herds were larger than 20, the cattle would be randomly selected and after sampling would be marked with blue dye to avoid re-sampling. Calves were only sampled over 1 years old.

The cattle were restrained by experienced handlers and livestock officers to reduce stress. Hair was collected using disposable razors to avoid contamination between animals. A sample of hair weighing > 0.04 g from the flank of each animal, and on roughly every third animal a sample was also taken from the belly (Figure 32). All hair was stored in heat-sealed aluminium bags.

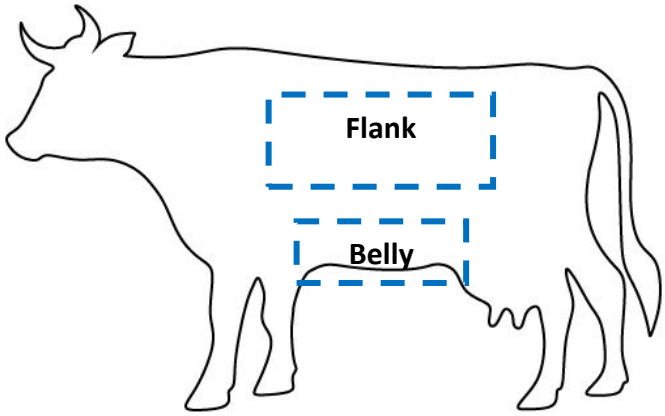

Figure 32: Diagram of the location where the flank and belly samples were taken from on the cattle.

3.2.1.3 Short-questionnaire

As each animal was sampled, the veterinarian asked the farmer specific information about any insecticide treatments which had been used. For each animal the data collected (Appendix A.4) included the date it was last treated, how the insecticide was applied (sprayed/dipped/pour-on), whether it had been entirely covered or if selective application had been used and the brand of insecticide used.

3.2.2 Methods for Tanga sampling and bioassays

3.2.2.1 Hair collection

To optimise the IQK, hair was analysed which contained a known concentration of insecticide at known times post treatment. To produce these samples, hair samples were taken from cattle belonging to the Tanzanian Vector and Vector-Borne Disease Research Institute (VVBDRI) at their ranch in Mivumoni in Tanga region in August - September 2017. Cattle at the Mivumoni ranch are currently treated weekly in a dip with amitraz.

Samples of hair were taken 3 days before spraying began to gain baseline samples from 8 cows. Tail hair was cut to aid with identification. On day 0, Amitraz (Tixfix, Rotam) was applied to four cattle at a concentration of 30 mL per 15 L and a pyrethroid (Paranex, Farmbase Limited) was applied to four cattle at a concentration of 8 mL per 15 L (Figure 33). Both insecticides were applied using different knapsack sprayers. The four cattle treated with Paranex were separated from their herd so as not to contaminate the four cattle treated with amitraz which could cause false results. The cattle were sprayed in a crush and the crush washed down afterwards. When taking hair samples the amitraz

treated cattle were placed in the crush before the Paranex treated cattle to reduce transfer of pyrethroid between cattle.

Figure 33. Photograph shows VVBDRI cattle being treated with insecticide on day 0.

Samples were then taken on days 1, 3 and 7 for both treatment groups. Samples were also taken on day 14 for the Paranex treated cattle only. Manufacturers guidelines state that amitraz must be reapplied weekly and Paranex every two weeks. The amitraz group could not be left for more than 7 days due to the risk of tick-borne disease.

3.2.2.2 Bioassays

Flies from the colony at the VVBDRI in Tanga were used to test the efficacy of the insecticides applied to the study cattle. Approximately 0.05 g of hair was placed into the top of a piece of pipe (~20 mm diameter), lined with parafilm. Parafilm was used so it could be replaced at the end of each test. Individual tsetse (*G. pallidipes*) were placed into 50 mL falcon tubes (25 mm diameter) within the colony building and transported to a separate laboratory on the edge of the compound, to reduce the risk of contamination from the hair samples to the main colony. Each tube containing a fly was placed on top of the pipe and pressed down so that the fly was forced to come into contact with the hair sample. The tubes were left in place for 45 s (Figure 34) before being removed (Torr, 1985), recapped and placed into a cool box containing a damp cloth for 2 h.

Figure 34. Photos show bioassay set up. Tubes were pulled down to ensure tsetse would remain in contact with the hair samples for the full 45 sec. Separate boards were used for each treatment and the controls.

Around 60 flies were used to test each treatment for each sampling day – being split into groups of approximately 15 flies for each cow. Around 30 flies were also used in a control group which were placed onto piping which contained no hair for each round of testing.

After two h the tubes were removed from the coolbox and examined to determine the effect on the flies. Flies would be graded as alive, moribund-front, moribund-back or knocked down according to whether they appeared (i) unaffected, (ii) affected but standing, (iii) affected and lying on their back or (iv) on their back and completely motionless. Tubes were re-used between experiments but were washed repeatedly.

To assess the effect of different concentrations of Paranex on tsetse, hair taken from the tail of untreated cattle was treated with Paranex applied at the manufacturer's recommended concentration (0.5 mL/l) or at 10% or 1% of the recommended concentration and bioassays conducted.

3.2.3 Insecticide analysis

3.2.3.1 Insecticide extraction

Hair samples collected from the cattle sampled in Serengeti and Mivumoni Ranch were returned to LSTM for extraction and analyses. Four samples from each herd were randomly selected as a sub-sample for GCMS analysis. A small part of each sample (0.05 - 0.15 g) was removed from the aluminium bag they were stored in and then weighed. Each sub-sample was then transferred to an individual glass tube, and 4 mL of acetone added on top of the hair sample, ensuring the whole sample was covered. The glass tubes were then sonicated

for 15 min. The 4 mL of acetone was then removed from the tube and pipetted into a fresh tube. Another 4 mL of acetone was placed onto the same hair sample to undergo another wash step, before sonicating again for 15 min. The 4 mL from the tube containing the hair was then removed and placed into the tube containing the 4 mL from the previous sonication step, giving 8 mL of acetone containing the insecticide, with no hair.

The samples were divided in half so that half could be retained at LSTM for analyses by HPLC and/or IQK and half sent to the Natural Resources Institute (NRI) for analysis by GCMS. Due to the high volatility of acetone, it was not possible to know how much acetone had evaporated during the time taken to sonicate. The 8 mL of solution was therefore evaporated and then resuspended in 4mL of acetone. This 4 mL was then quickly split into two tubes to ensure the sample was split in half as accurately as possible, before both sets were evaporated again. The set for NRI was then posted and the set for LSTM placed into a freezer (-20 °C).

3.2.3.2 GCMS

All of the bioassay samples, and a subset of 176 Serengeti samples were sent to NRI for analysis. Methods provided by NRI for GCMS are as follows. Samples were dissolved in 1 mL acetone (Pesticide Residue Grade, Fisher) and analysed by GCMS on an Agilent 6890 GC coupled directly to an Agilent 5973 Mass Selective Detector (Agilent Technologies, Manchester, UK). The GC was fitted with a fused silica capillary column coated with DB5 (30 m x 0.125 mm i.d. x 0.125 µm film thickness; Agilent). Carrier gas was helium (1 mL/min), injection was splitless (300 °C), and the oven temperature was programmed from 60 °C for 2 min then at 10 °C/min to 300 °C and held for 10 min. The transfer line temperature was 250 °C, quadrupole 150 °C and ion source 230 °C.

Injections of 2 or 3 µl were made with an autosampler, and total ion current (TIC) monitored from *m/z* 30-450. Cypermethrin and alpha-cypermethrin were detected by single ion scanning at *m/z* 163 and 181 and verification of the mass spectrum where possible. Calibration curves of cypermethrin were run from 0.3-30 ng/µl and quantification was done by these external standards using *m/z* 163.

Cypermethrin gave four approximately equal-sized peaks (8 pairs of enantiomers) with peaks 1, 2 and 3 resolved but peaks 3 and 4 only partially resolved. Alpha-cypermethrin showed generally only peaks 1 and 3 with peak 3 predominating, corresponding to the two *cis* isomers, although in some samples peak 2 was also present. In some samples, tetracosane (24:H; 50 ng) was added as internal standard and quantification used the TIC

for 24:H and m/z 163 for the cypermethrin, but this was found to be less reliable. GCMS data was then analysed using linear models in R for a range of variables.

3.2.3.3 HPLC

The evaporated bioassay samples were re-suspended in 1 mL of acetone and vortexed for 1 min. From this, 200 µl of each sample was removed and placed into a clean tube. Both the original sample and the sub-sample were then re-evaporated and the original sample returned to the freezer. The 200 µl samples were then re-suspended in 200 µl of acetonitrile. This was then filtered through a 0.2 µm syringe filter and 80 µl pipetted into a HPLC vial. A standard curve was also created using stock alpha-cypermethrin in acetonitrile from 1.9 µg/mL to 1000 µg/mL and a blank containing only acetonitrile.

The standards and samples were analysed using an Agilent 1100 series HPLC. The quantity of each sample injected was 20 µl with a 24 min run at 23 °C. The monitoring absorbance was 232 nm through a C18 column with a flow rate of 1 mL/min with an isocratic mobile phase of 70% acetonitrile and 30% water (Yunta *et al.*, 2016). The retention time for alpha-cypermethrin was 21 min.

The outputs from the analyser included the area under the peak produced. For the standards this area was then plotted against the known concentrations. The equation from the line of best fit was then used with the peak areas of the bioassay samples to calculate the concentrations.

3.2.3.4 IQK

The IQK comprises Reagents A, B and C. Reagent A releases cyanide ions, Reagent B is used for detection of the cyanide ions and causes the colour change reaction, and Reagent C stops the reaction. Reagent A was made by dissolving 75 mg of potassium hydroxide in 5 mL of water, then 95 mL of 95% ethanol was added to give a solution of 0.075% KOH in 90% ethanol. Reagent B was made by dissolving 400 mg of 2, 3, 5-triphenyltetrazolium chloride (TTC) in 5 mL of distilled water to which was added 40 mg of 4-nitrobenzaldehyde (PNB) dissolved in 95 mL of 95% ethanol. Reagent B was placed in the dark due to light sensitive components. For reagent C, 0.5 mL of glacial acetic acid was diluted in 99.5 mL of distilled water.

A set of standards were then made for the IQK ranging from 0.125 µg/mL to 15 µg/mL of insecticide in each tube. A solution of alphacypermethrin diluted in acetonitrile at a concentration of 10 mg/mL of alphacypermethrin was used to produce solutions of 1 mg/mL, 10 µg/mL and 1 µg/mL. The 1 µg/mL stock was used to make standards from 0.125

µg/mL to 1.25 µg/mL, and the 10 µg/mL stock was used to make standards from 2.5 µg/mL to 15 µg/mL.

Three sets of standards were made up so that a time series could be performed. The samples were run for a total of five, ten and fifteen min with the same concentrations. For each tube, 0.8 mL of Reagent A was added and vortexed, followed by 0.8 mL of Reagent B. They were then vortexed again and left for their designated time, before 0.4 mL of Reagent C was added and vortexed again. A photograph was then taken of each group with a negative control which contained only the reagents. For the standards, 200 µl of each was taken and placed into a microplate. For the samples collected from cattle in Serengeti district, 200 µl of each was also taken and placed into the plate, but the rest of the sample was then filtered with a 0.2 µm filter before 200 µl of filtered sample was then placed into the plate as well.

The plate was placed into the plate reader (BioTek® Epoch Microplate Spectrophotometer), and the software Gen5 (BioTek® Gen5 Microplate Reader and Imager Software) used. The options selected on the computer were absorbance, endpoint, monochromators and wavelengths set at 480, 500 and 550 nm. Once absorbance values were exported, the value of the negative control was deducted from the value of each sample. These corrected values were then plotted against the known concentrations for the standards. The slope of this line was then used to calculate the concentration of the bioassay samples. The values were then adjusted to account for the 20% of the original sample which had been removed for analyses by HPLC by multiplying by 1.25.

3.3 Results

3.3.1 Analysis of cattle sampling data

The centre of each village was within 5 km of the border of the protected area, and the households themselves no further than 10 km. A total of 44 herds were sampled and 44 questionnaires completed. The total number of cattle belonging to these herds was 1971 and of these, 958 hair samples were collected from 721 individual cattle.

All farmers reported that their entire herd had been treated, rather than selecting a small group of animals for treatment. All farmers who had sprayed their cattle reported that they sprayed the animals all over their body, and did not use a restricted application protocol (Torr, Maudlin and Vale, 2007).

Of 44 herds sampled, 68% (30/44) of herds had either been sprayed or dipped within the previous month, and 36% (16/44) within the previous two weeks. Of 27 sprayed herds, 70% (19/27) had been treated within the previous month and 48% (13/27) within two weeks. For the 17 dipped herds, 64% (11/17) had been treated within the previous month, and 18% (3/17) within two weeks. Of those herds treated within the last month (n=30), there was a wide range in days since the herd had last been treated, but the average days since treatment for spraying was 11 and for dipping was 15 (Table 3). The average herd size of those which were sprayed was 55 (range 4 – 280), and the average herd size of those which were dipped was 29 (range 11 – 57).

Table 3. Of cattle treated within the previous month of the sample collection date, the table shows the average times since cattle had been treated with insecticide.

Treatment type	Average time since last treatment (d)(n)	Range (d)
Spraying	11 (19)	1-26
Dipping	15 (11)	2-20

The mostly commonly used insecticide formulation was Albadip followed by Paranex. Alpha-cypermethrins were used on 75% of all herds, and cypermethrins used on 20% of herds (Table 4).

Table 4. Table shows the last type of insecticide used by farmers in the study, and the percentage of herds which had used each.

Insecticide	Insecticide type	Percent of herds used on	N
Albadip	Alpha-cypermethrin	50%	22
Paranex	Alpha-cypermethrin	25%	11
Cybadip	Cypermethrin	16%	7
Tantix	Cypermethrin	4.5%	2
Farmer unsure	NA	4.5%	2

Compared to the overall percentage of cattle treated with Albadip for both sprayed and dipped herds (50%), Albadip is less commonly used in herds which have been spraying (24%), as Paranex was the most commonly used (41%) (Table 5).

Table 5. The last type of insecticide used by farmers who had sprayed their cattle.

Insecticide	Insecticide type	% of herds used on	N
Albadip	Alpha-cypermethrin	26%	7
Paranex	Alpha-cypermethrin	41%	11
Cybadip	Cypermethrin	19%	5
Tantix	Cypermethrin	7%	2
Farmer unsure	NA	7%	2

For herds which had last been dipped, Albadip was the insecticide used by the majority of farmers (Table 6), which is much higher than the overall percentage of herds which used Albadip (50%).

Table 6. The last type of insecticide used by farmers who had dipped their cattle.

Insecticide	Insecticide type	% of herds used on	N
Albadip	Alpha-cypermethrin	88%	15
Cybadip	Cypermethrin	12	2

The use of insecticide product was not spread evenly across villages (Figure 35). In Bwitengi, Makundusi, Robanda and Singisi, Albadip was the most commonly used insecticide whereas at Bisarara and Park Nyigoti the percentage of brands used was split more evenly and four types were used.

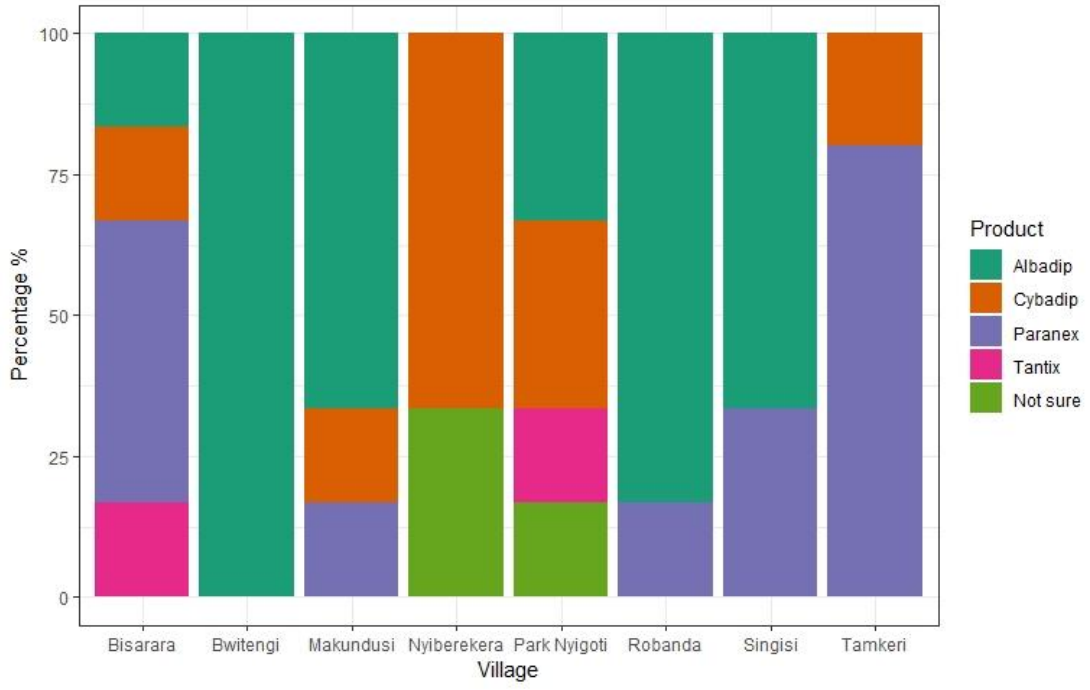

Figure 35. Percentage of each insecticide product used within each village. All villages had 6 herds questioned apart from Tamkeri (n=5) and Nyiberekera (n=3).

In four of the eight villages, the last reported method of treatment was the same across the whole village (Table 7). In Bisarara, Nyiberekera and Tamkeri, all households were applying insecticide through spraying, and in Bwitengi all households were using a dip. The other four were divided between spraying and dipping, with Makundusi and Singisi having a higher percentage of households using spraying, Robanda having a higher percentage of households dipping, and a 50:50 split at Park Nyigoti.

Table 7. Method used to treat cattle with insecticide within each village

	Spraying (%)	Dipping (%)	Total herds
Bisarara	100	0	6
Bwitengi	0	100	6
Makundusi	66.667	33.333	6
Nyiberekera	100	0	3
Park Nyigoti	50	50	6
Robanda	16.667	83.333	6
Singisi	83.333	16.667	6
Tamkeri	100	0	5

3.3.2 GCMS analyses of cattle hair

The GCMS analyses showed that 18% (31/176) of individual cattle and 27% (12/44) of herds had detectable levels of cypermethrin. Only 11 herd (Figure 36) had detectable levels of insecticide, however one herd (CS-BW02) also had detectable but un-quantifiable amounts of insecticide present. The range of detected insecticide amounts was from 1.6 µg/g to 1278 µg/g. The mean concentration from only the samples where insecticide was present was 194 µg/g, and when taken from all samples the mean was 30 µg/g. When analysed by herd, there is often a lot of variation within herds which had positive samples. Only 33% (4/12) of the herds with positive samples had insecticide detected on all four of their samples. For herds that had 3 out of 4 positive samples, this was 17% (2/12), for 2 out of 4, this was 25% (3/12), and for herds with only one out of four samples positive this was also 25% (3/12). Within villages, three villages had more than one herd with samples positive for insecticide (Makundusi, Singisi and Bisarara) and three villages had only one herd with samples positive for insecticide (Robanda, Bwitengi and Park Nyigoti).

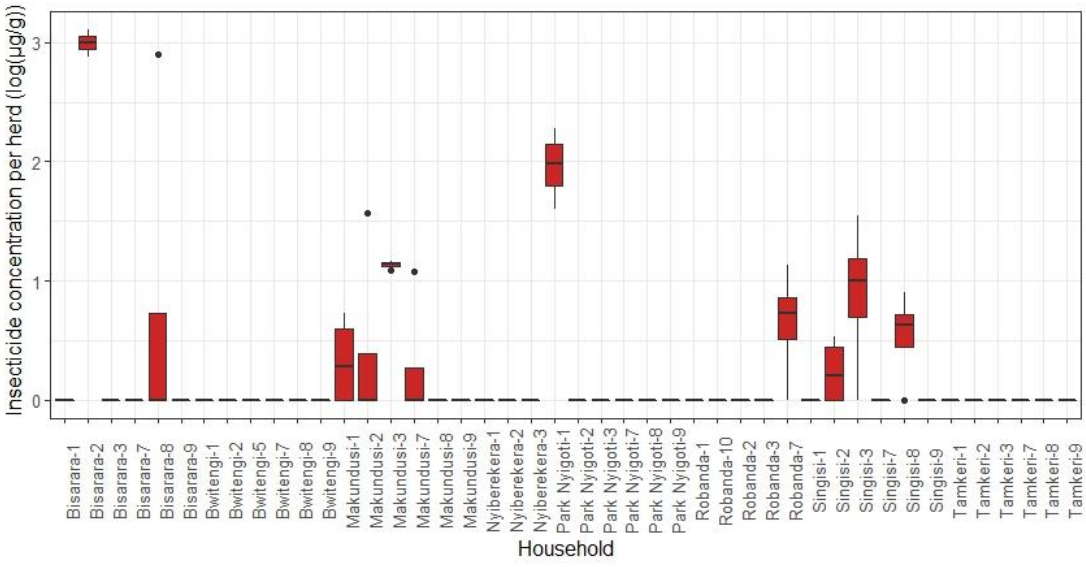

Figure 36. Insecticide concentration by household

Out of eight villages, six villages had cattle where insecticide was detected. Bwitengi is not shown as having insecticide present in Figure 37, however small amounts were detected which could not be quantified. The highest concentrations were found at Bisarara and Park Nyigoti. No samples were positive for insecticide at Nyiberekera or Tamkeri.

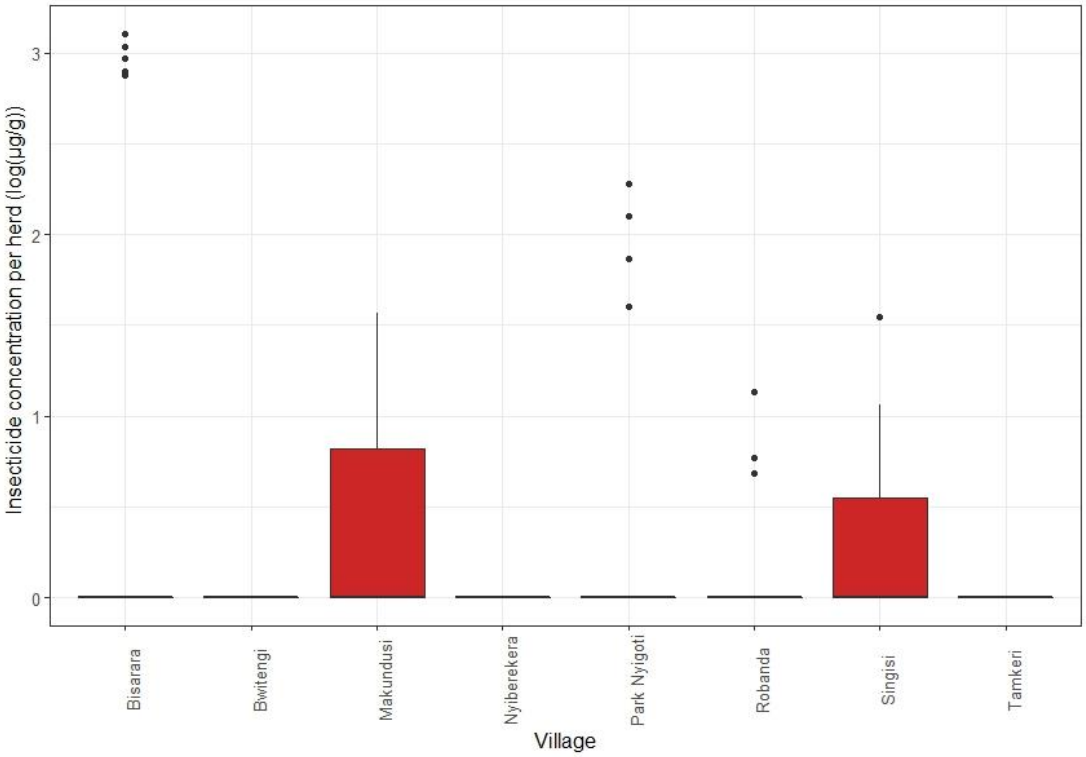

Figure 37. Insecticide concentrations by village.

A linear model for the log insecticide concentration compared with the log of the days since the farmer claimed they had been treated was statistically significant with a P value of 0.00272 and an adj. R^2 value of 0.1755. The number of cattle in a herd was not significantly related to the insecticide concentrations (Figure 38) found with a p value of 0.468 and an adj. R^2 value of -0.01088. The distance to a protected border was not significantly related to insecticide correlation (Figure 38) with a P value of 0.34 and adj. R^2 value of -0.001835. The age of cattle was not significantly correlated with insecticide concentration (Figure 38) either with a P value of 0.194913 and adj. R^2 value of 0.004, however there is a general trend that as cows get older past 3 years, the highest amount of insecticide decreases, although a small sample size for cows older than 6 years reduces the statistical power. The condition score (estimated energy reserves in the form of fat and muscle) of the cattle was not significantly correlated with insecticide concentration, nor was the sex of the cattle.

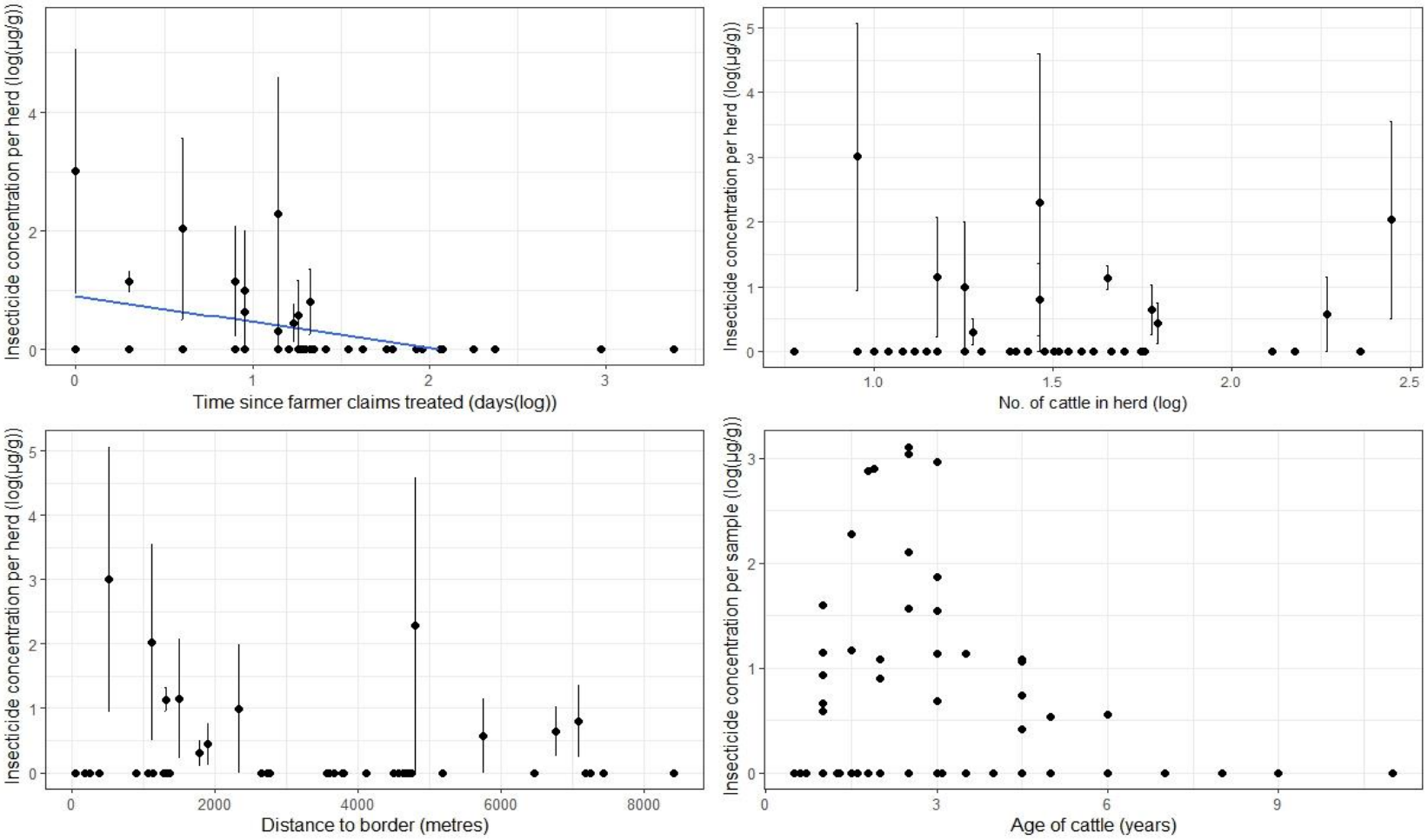

Figure 38. Insecticide concentration compared with time since the farmer claims to have treated the herd (TL), with the number of cattle in a herd (TR), with distance of the herd to the border of a protected area (LL), and with the age of the cattle the sample was taken from (LR). Error bars show standard error of the mean and the blue trend line indicates a significant relationship.

3.3.3 Bioassays

The percentage mortality of tsetse exposed to hair samples treated with Paranex declined gradually over the two-week sampling period. The day after treatment, mortality was 100% and flies exposed to the hair were knocked down soon after exposure to the treated hair. At seven days post exposure, knockdown was still relatively high at 63.6%, whereas at 14 days after exposure knockdown was almost 0% (Figure 39). Flies exposed to hair treated with amitraz were not knocked down at any point in the experiment. None of the flies exposed to untreated hair were knocked down. For tsetse exposed to Paranex-treated hair, the percentage of tsetse which were both moribund and knocked down at the end of the two h was 100% for hair collected within a week of treatment. This dropped to 24% by the end of the second week (Figure 6). None of the flies exposed to the amitraz treated hair were found to be moribund at any point in the experiment.

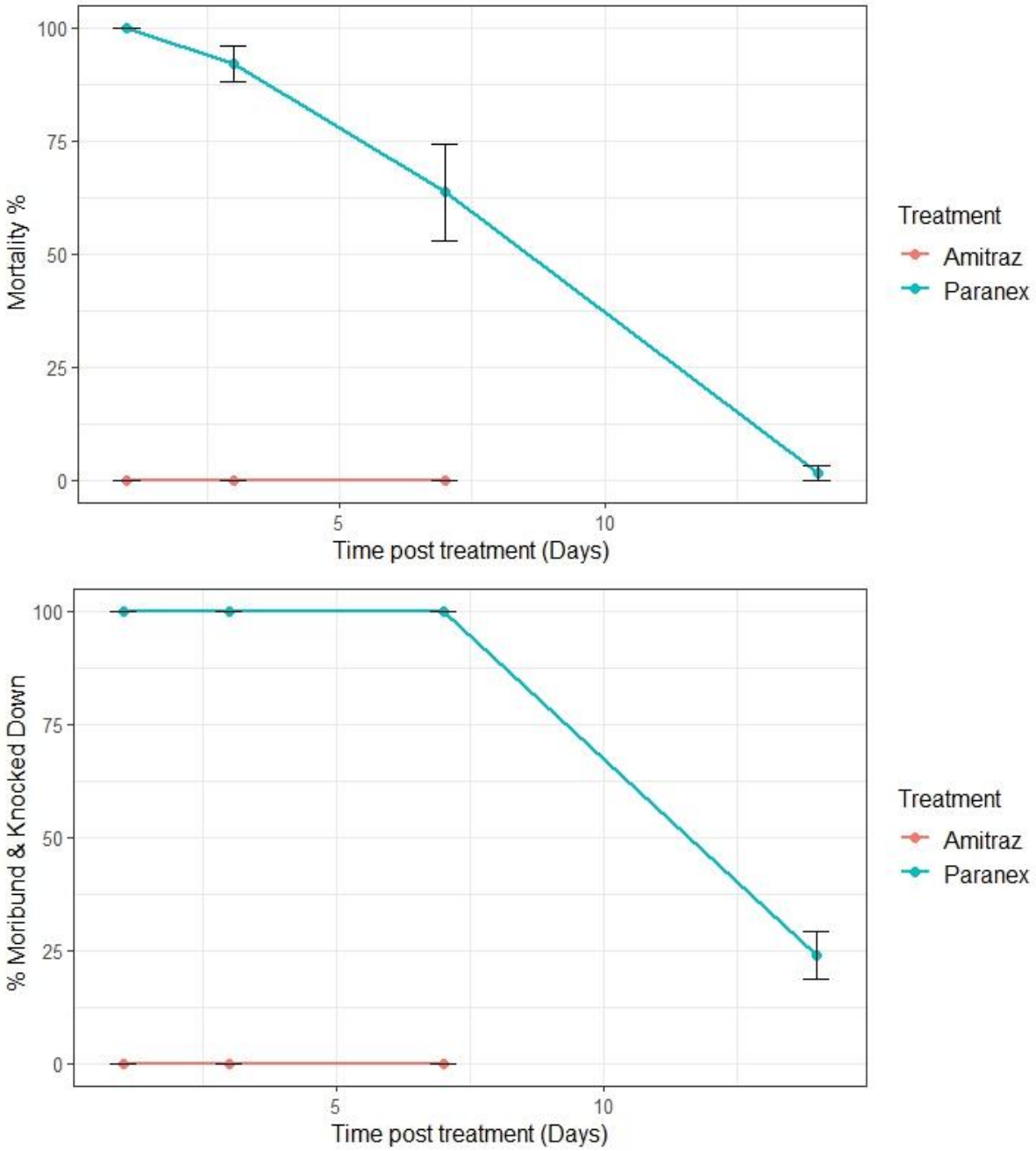

Figure 39. The percentage of tsetse which were knocked down (Top) and moribund and knocked down (Lower) at the end of the two h observation period after being exposed to hair samples which were collected from cattle on days 1, 3, 7 and 14 post treatment. The percentage mortality was taken as a mean from across the four cows used for each treatment. Error bars show standard error of the mean.

For tail hair that was treated with Paranex in the laboratory, the percentage mortality dropped by 21% when the insecticide concentration had been reduced by 90%. The percentage of flies which were either knocked down or moribund did not decrease between 100% and 10% of the recommended concentration (Figure 40). Even at 1% of the recommended concentration, almost all of the flies were either moribund or knocked down

so would not have been likely to survive much longer after the two h time period used to measure effect.

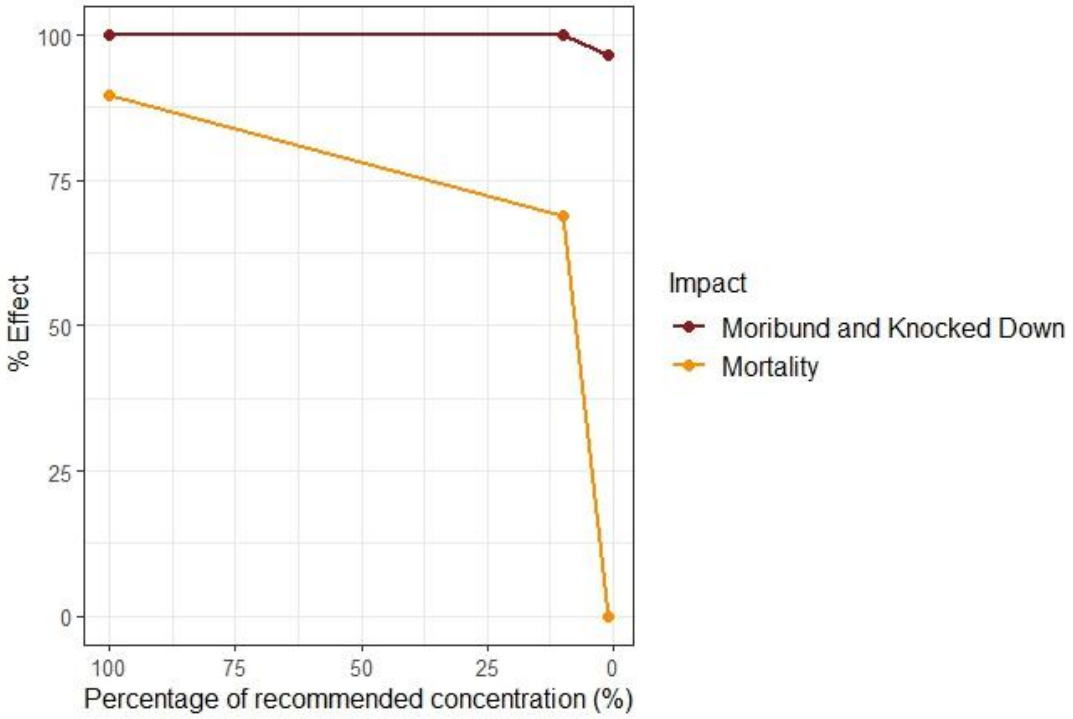

Figure 40. Percentage of flies which were either knocked down, or knocked down and moribund after exposure to tail hair treated with Paranex at the manufacturers recommended concentration (100%), at 10% of the recommended amount and at 1%.

The mean insecticide concentration of insecticide detected by GCMS from samples collected on day 1 post-treatment was 453.4 µg/g (Figure 41). By day 14 post-treatment, insecticide was no longer detectable. The concentration dropped most steeply between days 1 and 3, less so between days 3 and 7 and the least between days 7 and 14. The standard error was greatest in samples collected on day 3 post-treatment but decreased to levels comparable with day 1 by day 7.

There is a non-linear relationship between insecticide concentration on the bioassay hair samples and mortality of the flies exposed to them (Figure 41). In the first week, mortality drops by 36% while the insecticide concentration has dropped by 85%. Between days 7 and 14 post treatment, mortality decreases by 62% while insecticide concentration drops by 15%.

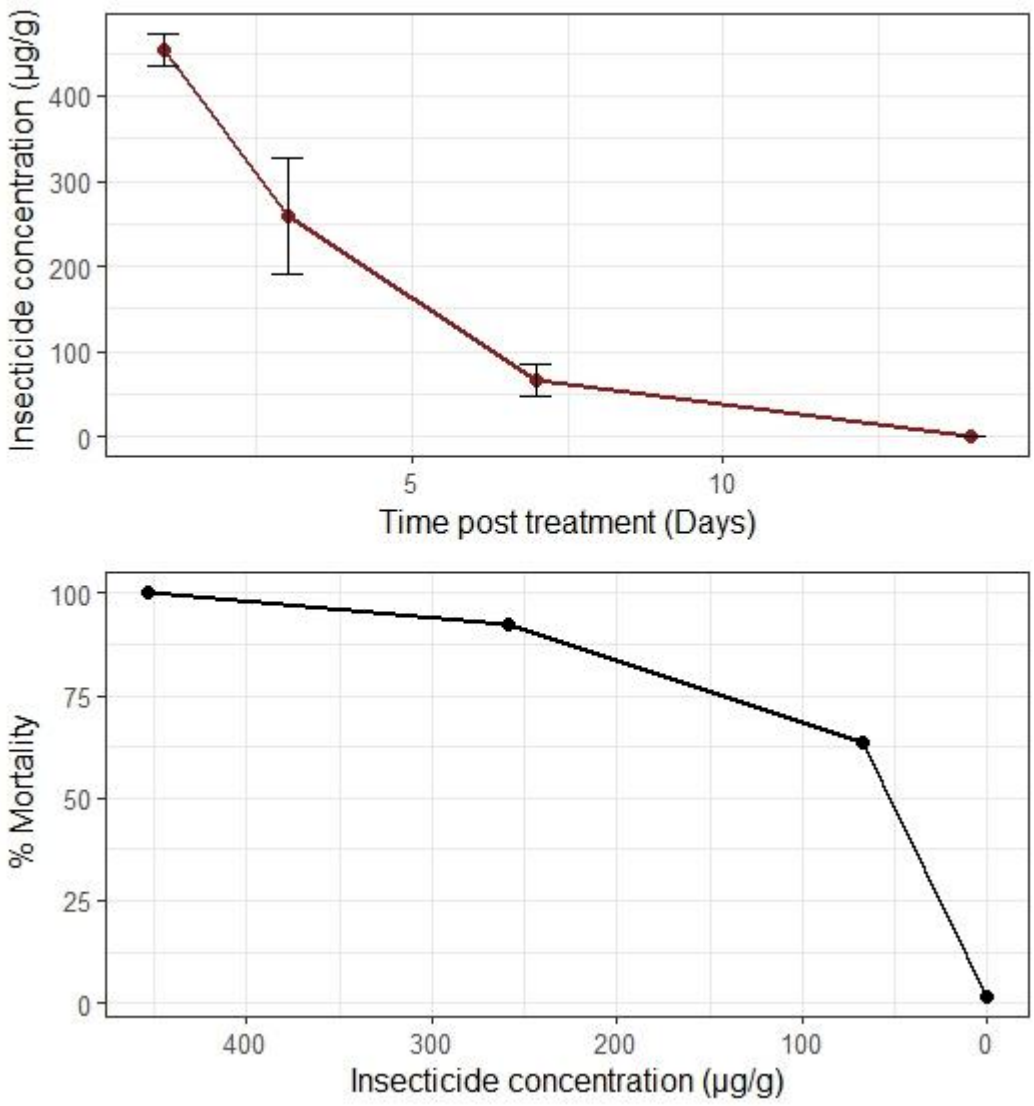

Figure 41. Insecticide concentration of bioassay samples over time averaged over the four cattle with error bars showing standard error of the mean (Top), and the relationship between insecticide concentration on the bioassay samples and the percentage mortality of the flies exposed to them. The furthest left point shows the data from samples collected on day 1 post-treatment, and the furthest right shows data from samples collected on day 14 post-treatment (Lower).

3.3.4 HPLC standards

Standards for HPLC were made up with a range of amounts from 0.244 µg/mL to 1000 µg/mL in each tube. The standards for 0.244 µg/mL and 0.488 µg/mL were undetectable on HPLC, however the rest of the samples were detectable and had an R^2 value of 0.999 (Appendix A.1) showing a highly related relationship between concentration and the area under the peak detected. The slope of the line was then used with the peak area of the

bioassay samples to calculate the insecticide concentration. Those values were then multiplied by five as they were only 20% of the initial sample.

3.3.5 HPLC vs GCMS

The relationship between the GCMS values for the bioassay samples was compared with the values gained by HPLC (Figure 42). A linear model comparing the two gave an adj. R^2 value of 0.948 with a P value of 2.74^{-9} which was highly significant. Although the fit is good, the actual amounts detected are generally lower by HPLC.

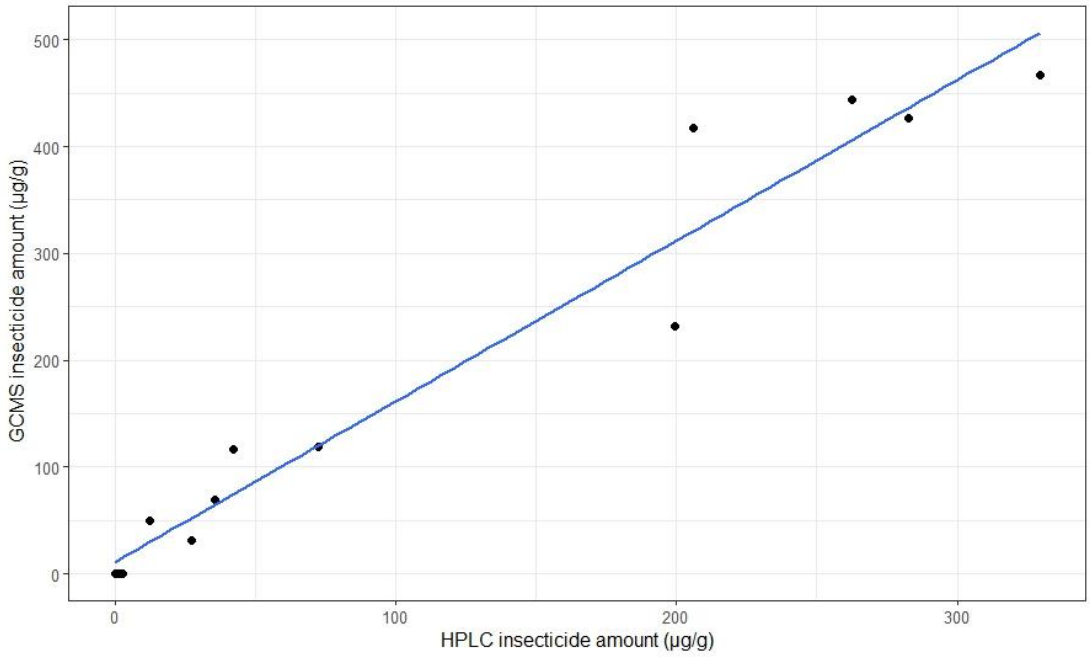

Figure 42. Insecticide concentrations of bioassay samples measured with both GCMS and HPLC.

3.3.6 IQK standards

To develop the IQK a set of standards were measured for their absorbance values across a likely range of concentrations based on the GCMS and HPLC results. Standards up to 15 µg/mL were tested however after 15 min the concentrations over 10 µg/mL had begun to saturate (Figure 43). The R^2 values for the 5, 10 and 15 min standard curves were 0.967, 0.9849 and 0.9373 respectively.

As the samples had begun to saturate, the curves were replotted with 10 µg/mL as the highest concentration. The R^2 values for the 5, 10 and 15 min standards were 0.996, 0.9992 and 0.9986 respectively (Figure 43) which were all highly related. There was a very slight background colour in the negative controls, therefore the absorbance value of these was subtracted from the absorbance of the samples with insecticide in. The 15 min test was selected to attempt to give more visual differentiation at lower concentrations (Figure 44),

and the slope from this line was used to calculate the concentrations of the bioassay samples based on their absorbance.

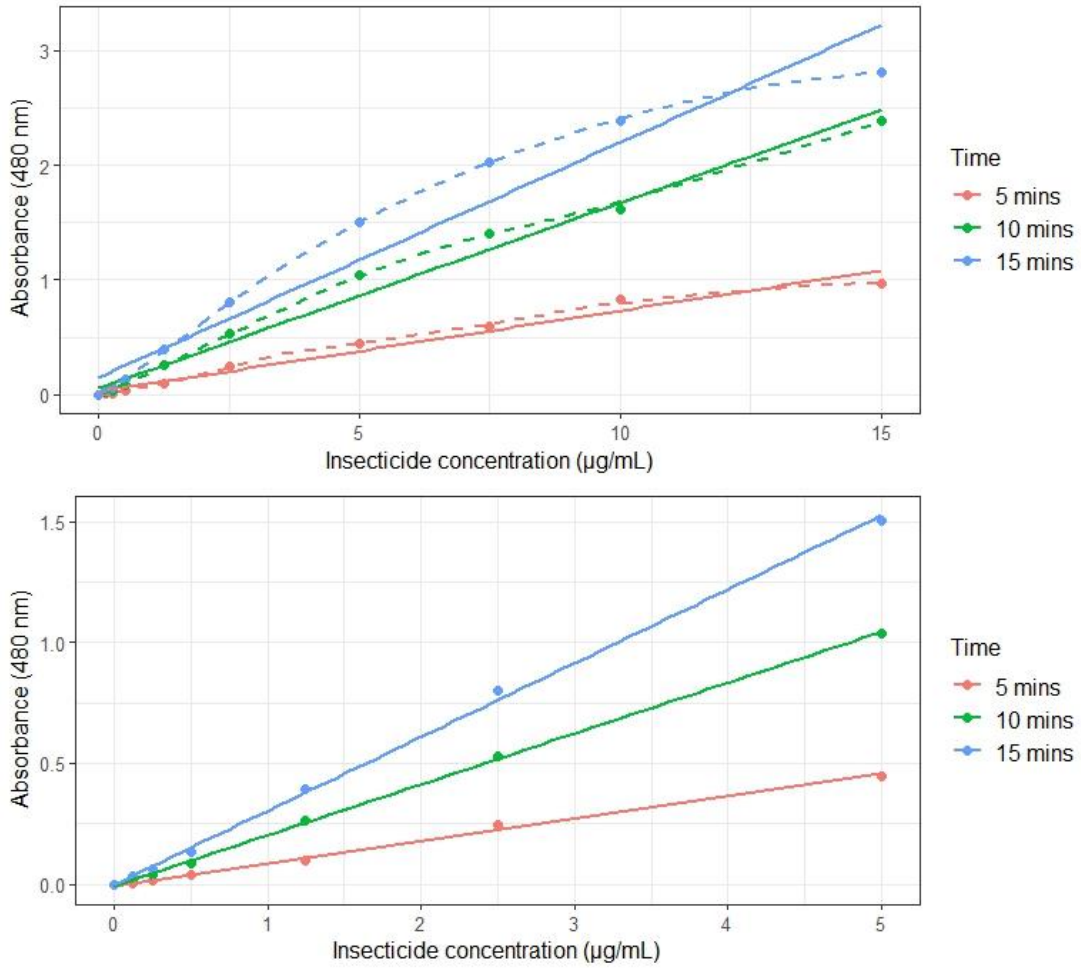

Figure 43. The solid lines show a linear relationship between the points and the dashed line is following the points to show the saturation in the 15 min curve (Top), and the standards when only up to 5 µg/mL (Lower).

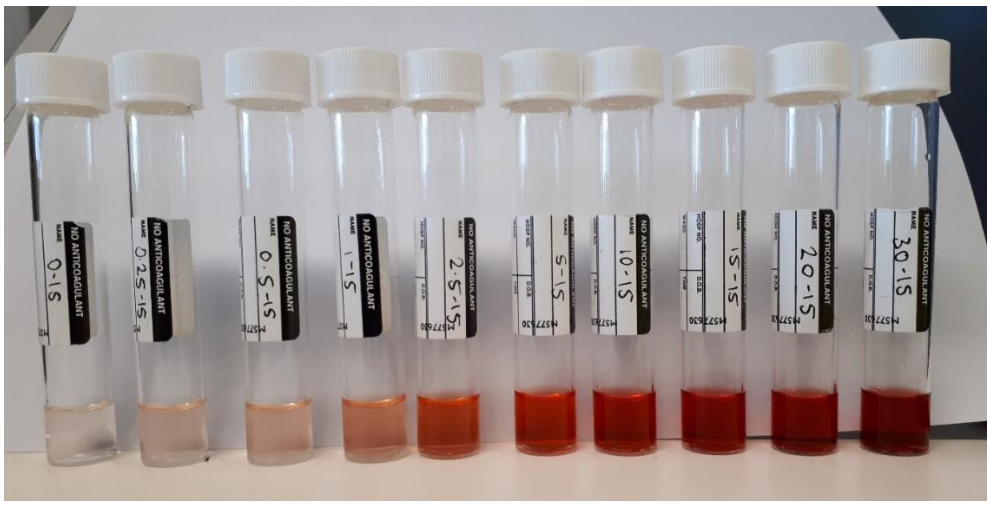

Figure 44. Photograph of the alpha-cypermethrin standards measured with the IQK at 15 min.

3.3.7 IQK on bioassay samples

Ideally, the IQK would be optimised so that it could be used in the field without a colourimeter to provide clear and obvious results. From visual inspection, the samples from days 1 – 3 post treatment (Figure 45) are clearly red in comparison to the negative control. The samples from days 7- 14 (Figure 45) are slightly more difficult to distinguish from samples that may just have dirt in them, although a light pink colour is still visible in the day 7 samples. These photos were taken before filtering and afterwards it was more obvious as a pink rather than a brown colour.

From visual inspection, the samples taken three days before treatment were much paler (Figure 45) than the Paranex treated samples. There is slight colouration in several of the samples however this looks much more brown and likely to be from dirt, and after filtering the samples became much clearer without becoming pinker. Overall, there was a clear difference between the day 1 and day 14 samples, however it was less clear between days 1 and 3, or days 7 and 14.

The samples which had been treated with Amitraz looked slightly darker (Figure 45) than those from before treatment began, but it appears to just be a darker brown, and they were noticeably dirtier than the pre-treatment samples before IQK reagents were added to the tube. After filtration the samples were much clearer.

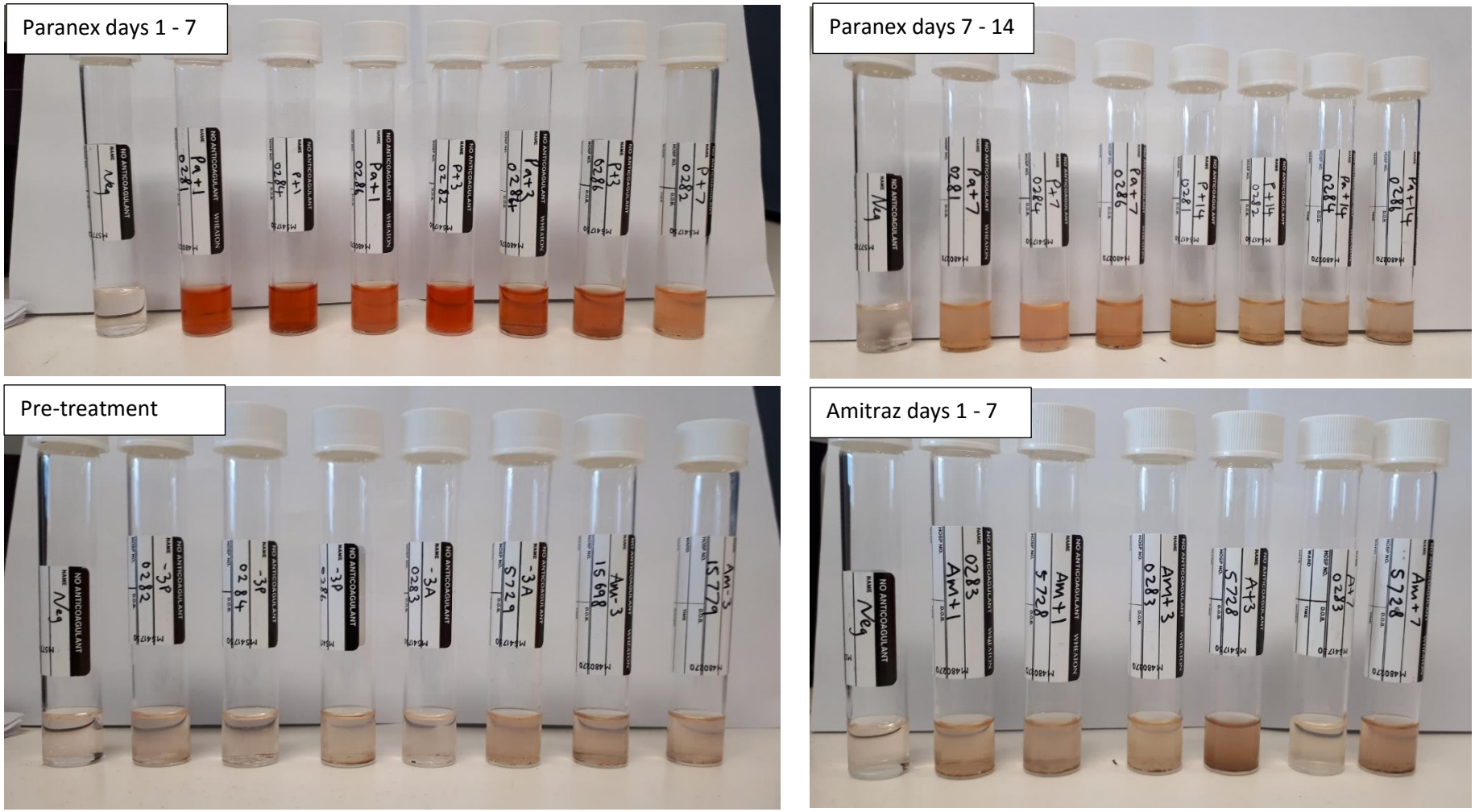

Figure 45. Photograph of IQK reactions from several of the bioassay samples from days 1 to 7 post treatment with Paranex (TL), samples from days 7 to 14 post treatment with Paranex (TR), samples taken three days before treatment with began (LL, and samples taken on days 1 – 7 post treatment with Amitraz (LR).

3.3.8 Method comparisons

The absorbance values from the Paranex treated samples were converted into concentrations using the slope from the 15 min standard and compared with the concentration values gained from GCMS (Figure 46). The actual difference between concentrations calculated through IQK and GCMS were in most cases greater than the difference between the IQK and HPLC values however a very similar trend was seen. The IQK samples with the higher concentrations were also higher by GCMS, with an adj. R^2 value of 0.854 with a P value of 1.9^{-7}.

The IQK values were all lower than the HPLC values for each sample, however the relationship between the two was significant (P: 8.12^{-11}) and had an adj. R^2 value of 0.971. The results from both tests were positively correlated with each other.

The average multiplication factor difference between IQK values and GCMS values from days 1 – 7 was 14.9 (range 3.6 – 59.5). The samples from day 14 post-treatment were discounted as they showed absorbance values similar to the background readings. The adj. R^2 value when re-plotted was 0.776 with a P value of 9.48^{-5} which was less than for the relationship of GCMS with the original IQK value. The average multiplication factor difference between IQK values and HPLC values from days 1 – 7 was 8.1 (range 2.8 – 21.4). The adj. R^2 value when re-plotted was 0.958 with a P value of 5.64^{-7} which was less than for the relationship of HPLC with the original IQK value.

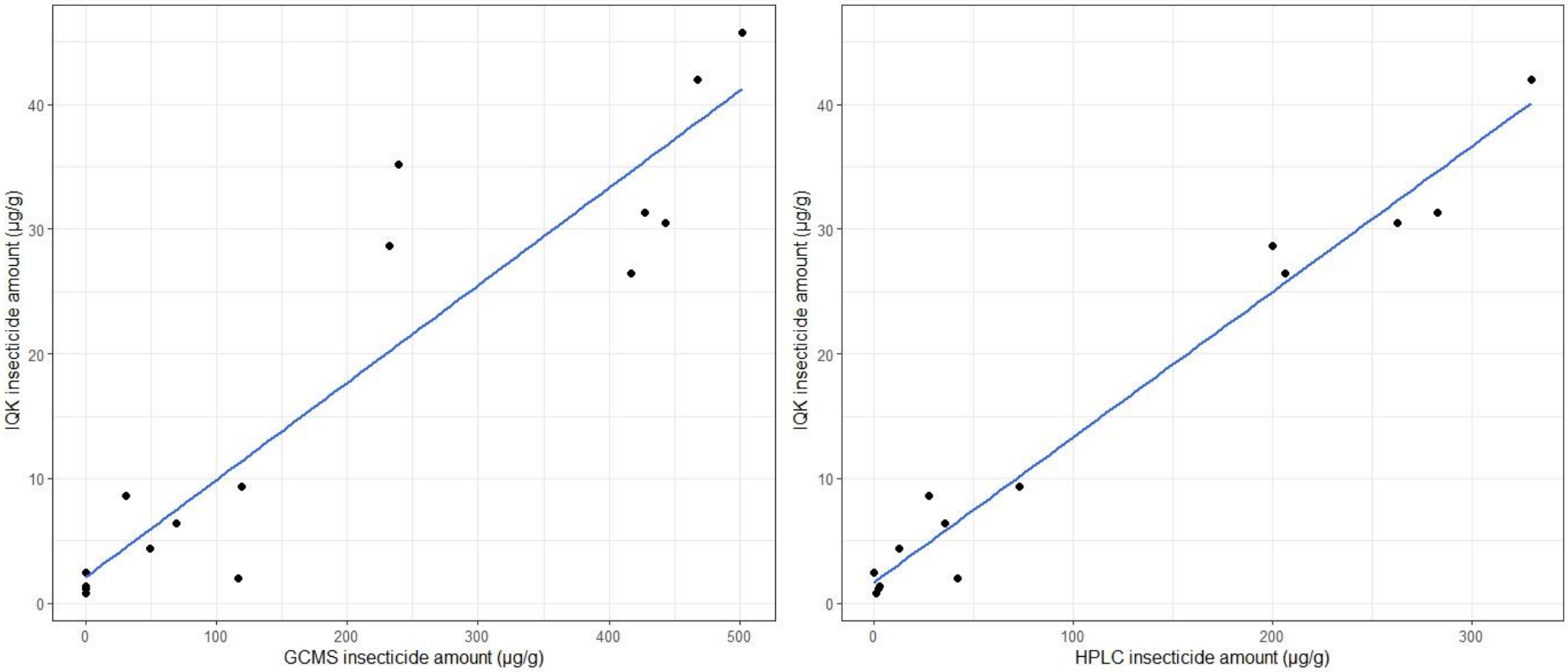

Figure 46. Bioassay samples measured with both the IQK and GCMS (L), bioassay samples measured with both the IQK and HPLC (R).

3.4 Discussion

3.4.1 Serengeti sampling

Uptake of ITC as a method of vector control appears to be high across the cross-sectional study area. These herds were recruited randomly to ensure no bias towards farmers which had self-identified diseases in their livestock to researchers as they had in the ZELS longitudinal study. Despite this, 36% of farmers claimed to have applied insecticide within the last two weeks, which is the recommended treatment interval for the formulations of alphacypermethrin being used. Treating a third of the cattle with a pyrethroid is likely have a significant impact on tsetse populations (Hargrove, 1988).

The average time since previous treatments for herds treated within a month means many of these animals will have been protected from tsetse. Although the average of 15 days for dipping was not as high as the average of 10 days for spraying, there are studies showing efficacy of insecticides within a month (Vale, Mutika and Lovemore, 1999). The average size of the sprayed herds was almost double that of the dipped, and a larger herd size generally indicates that the farmer is wealthier. Wealthier farmers may be able to afford to treat their herd more often which could also explain why the sprayed herds were treated more frequently. It may also be logistically easier to spray a herd than to take it to the dip tank, and it may be easier to take a smaller herd to a dip than it would be to take a large herd.

Products effective against only ticks were seen to be available in local veterinary supply shops and are often cheaper than those effective against ticks and tsetse per bottle. The fact that all of the insecticides reported to be used are effective against tsetse, likely indicates that government advice has successfully reached most farmers.

Robanda, Park Nyigoti and Bwitengi reported the highest use of dipping and these are three villages which had fully functioning community dips, as evidenced by a document (Appendix A.5) provided by the District Veterinary Officer in Serengeti District reporting the status of government dips in February 2017. In Bisarara, Nyiberekera and Tamkeri, all the households were spraying and there are no community dips listed in these villages which suggests that none of the farmers had their own dip tank or had access to any other privately-owned dips. The restricted spraying method had not been used by any farmers, and upon speaking with government officials, it transpired that the government had recommended treating animals all over their body due to the threat of tick-borne disease.

The bioassay samples showed that if insecticide had been applied following the manufacturers recommendation, alphacypermethrin could not be detected by GCMS 14 days after treatment. The number of farmers who said they had treated within 14 days was 36%, which is similar to the 27% of herds which had detectable levels of insecticide present. This was the only variable out of cattle age, distance of the herd from border, number of animals in the herd which significantly correlated with insecticide amount. Although farmers had stated that they had applied insecticide to their whole herd, it appears only some of them were actually doing this as only 33% of the herds which had positive samples, were positive for insecticide on all four samples which were tested. It is possible that the question had not been asked in the best way by the translator, or that the farmer had stated what they thought was the best answer. If farmers are only treating part of their herd, this would have implications on control, as the proportion of flies taking bloodmeals from treated cattle would be reduced and the mortality of the flies would be reduced (Hargrove and Williams, 1995), although evidence shows in some circumstances treating only part of the herd does not reduce the level of protection to the herd (Muhanguzi *et al.*, 2014).

3.4.2 Bioassay experiment

Flies exposed to Paranex treated hair collected on day 1 post-treatment were all knocked down within several minutes and was the same across flies exposed to hair from each of the four cattle. Although mortality had dropped to ~2% by day 14, approximately 25% of flies were left moribund so would not have been able to further transmit trypanosomes if they had taken an infected bloodmeal.

It was expected that no flies would be knocked down by hair treated with amitraz and this was confirmed. Bardosh, Waiswa and Welburn, (2013) discuss this effect but there is no published literature showing this. There is only literature stating that that it has no effect on the survival of mosquitoes exposed to it (Njoroge *et al.*, 2017). Despite the fact that the product does not state on the bottle that it is not useful against tsetse flies, there is little recorded literature which has described this, and this was not widely known by local support staff.

Although no insecticide was detected on hair samples from day 14 using GCMS, tsetse were knocked down on exposure to the samples. This suggests that insecticide was still present and effective, and the bioassay seems to be more sensitive than GCMS. The estimates of the numbers of cattle treated with insecticide from the Serengeti study is therefore likely to be conservative. There was however a time delay between the bioassay experiment and

the samples being extracted of several weeks. Whilst they were kept in a -20 °C freezer for the majority of that time, there were several days between travelling back from Tanzania to the UK when they were not kept frozen, and whilst in Tanzania were only kept in a fridge so the insecticide may have degraded.

The large standard error in GCMS values for samples from day 3 could be caused by a number of factors. It could be that the cattle had been exposed to different environmental or physical conditions over those days such as being more in direct sunlight or more rain, or they may have rubbed on the ground or against other cattle. It is also possible that there was an unequal application of insecticide across cows and samples may have been taken from areas where it had been well applied on one cow and sparingly on another, as by day 7 the standard error had reduced again.

The bioassay experiment using tail hair treated with dilutions of the manufacturers recommended concentration show that even if farmers were not making up their insecticides precisely, it is likely that it would still be very effective at killing tsetse, as even at 10% of the recommended amount, mortality has only dropped from 90% to 69%. This is reflected in Figure 41 by the non-linear relationship between the concentration of the bioassay samples and the mortality of the flies and the fact that mortality does not decrease at the same rate of insecticide concentration during the first week post-treatment. It was only after insecticide dropped to 85% of the original concentration that mortality began to steeply decrease.

The mean GCMS value for bioassay samples at day 7 was 66.8 µg/g which knocked down 63.65% but still left 100% of flies near death. Only 26% of the positive samples (8/31) had greater than 66.8 µg/g of insecticide detected on them, however even when the insecticide was undetectable it was still capable of leaving ~24% moribund, therefore it is likely that the moribund rate which would be produced by the remaining 23 positive samples would have been greater than 24%, ranging up to 100%.

3.4.3 IQK

When testing the standards for the IQK, the negative control appeared to have a slight colour change which developed over time. To ensure this was not due to contamination the standards were re-made using un-opened bottles of reagents and the same colour was seen. This colour change was sufficiently small enough that it did not overlap with the lowest concentration so was considered to not be a significant problem to the test and was most likely just a background colouration.

As GCMS did not detect any insecticide on days 14, the IQK appears to have a similar level of sensitivity. The lack of pink/red colouration in the pre-treatment and amitraz samples show that the test does not interact with components in the cattle hair such as oils or fats, and that it does not react with other chemicals. The figures comparing IQK with HPLC and GCMS (Figure 46) are on different axis scales. This is because there are two additional data points for GCMS which could not be gained for HPLC as the IQK had been run and 20% not already extracted as for the other bioassay samples. The R^2 value (0.9712) for the relationship between the IQK and HPLC was greater than the R^2 value (0.8536) for the IQK and GCMS. Whilst the data sets are slightly different, this could be due to the GCMS being less temperature stable than HPLC, and could have caused more variety in the results.

It appears as though the IQK is underestimating the insecticide concentration compared to both GCMS and HPLC, which could be a limitation of the extraction reagent in the IQK. Despite this apparent underestimation of the IQK, the very high recalculated R^2 values against HPLC (0.9575) and GCMS (0.7759), which are both reliable molecular methods, shows there is great potential for the IQK as a method for giving quantitative results for insecticide concentrations. If the IQK could be modified to be slightly more obvious at lower concentrations, it could eventually be used in the field. It could be used by farmers for testing the quality of the product they have bought, or for checking when insecticide needs to be reapplied. It could also be used by researchers to conduct quick testing of insecticide presence to gain an overview of use in an area or by technicians on national monitoring programmes, and would enable them to give farmers instant feedback, such as in the 'Stamp Out Sleeping Sickness' programme in Uganda (Bardosh, Waiswa and Welburn, 2013).

3.4.4. Limitations and solutions

Although there appears to be a difference between when farmers said they had been applying (68% claimed to have treated within a month and 36% within the previous two weeks) and what was detected with GCMS (insecticide detected on 18% of samples and 27% of herds) – this is likely due to the fact that the method was not able to confirm insecticide amounts comparable with 14 days post application if it had been applied correctly. It can be difficult to extract information from questionnaires if the interviewee is giving answers that they think the interviewer wants to hear. It is also sometimes difficult to get focused answers if their animals are being sampled at the same time as they may be concentrating on the welfare of their animals. As for the mini-questionnaire in this chapter it was focused on animal specific questions so this could not be avoided.

In terms of collecting the samples - shaving the cattle was not always easy if the hair was slightly longer than average or particularly dirty. It also created lots of waste as the razors were only used once. The extraction protocol is currently time-consuming and could also be refined.

To improve the hair sampling, a simpler way of collecting samples might be to cut a section of hair from the ends of the tail using scissors. Cattle sweep their tails repeatedly across their flanks and hence the tail hairs are likely to be contaminated with insecticide. As long as the scissors could be sufficiently cleaned between each sample to remove insecticide residue this would reduce waste and be a much quicker option. It would also need to be checked that it gave comparable amounts to be able to compare with the samples already collected from the Serengeti. Most tsetse feed on the lower legs of cattle (Torr, Maudlin and G. a. Vale, 2007) and the concentration of insecticide on this part of the body is the most important. Future studies should aim to quantify the concentrations in this area and determine whether a sample of tail hair provides a simple means of assessing whether an animal is treated with an adequate dose of insecticide.

For the extraction protocol, an experiment could be designed to test the efficacy of various extraction protocols – ranging from sampling placing the collected hair straight into the IQK reagent and vortexing or attempting it with only a single wash and sonication step compared to two, and then comparing the amount of insecticide that has been extracted. This would save many hours of work and would hopefully reduce the time the insecticide would spend in an unstable state potentially degrading. There are also hand-held colorimeters which could be tested to give even more accurate results in the field, as long as samples were first filtered. The Serengeti samples which were tested by GCMS will in the future also be tested by IQK once the extraction protocol has been refined. Ultimately, it should be possible to develop a simple field test which could be applied in the field to enable livestock officers to assess whether cattle have been treated with insecticides.

3.5 Conclusion

This study shows widespread use of pyrethroids at regular intervals within the sampled herds. While these results are only from 44 herds, if this is being replicated across the district and region it is likely that the use of insecticide has contributed to the reduction tsetse densities in areas bordering the Serengeti National Park and surrounding game reserves. While the reported and detected amounts of insecticide differ, both would have an impact on tsetse populations of the transmission of trypanosomes.

According to the literature (Randolph and Rogers, 1978; Hargrove, 1988), the proportion of herds treated in the Serengeti study area would be capable of reducing tsetse populations. From the evidence presented so far, it appears as though this could be a case study of where government policy has successfully impacted HAT disease risk, alongside farmers who have carried out much of the control on their own by trying to protect their cattle. Messages about not using restricted application have clearly filtered down so it is likely that other insecticide policies have been widely received. The percentage of farmers reporting insecticide treatment on a regular basis is very high and the reduced abundance of tsetse flies in farming areas indicates that this is likely to be having a significant impact.

This is a sign that a geographically-limited intervention by livestock keepers around a wildlife area has been successful. There has been warnings and concerns (Tanner *et al.*, 2015; Hartter *et al.*, 2016; Squarre *et al.*, 2016) raised about parks being sources of disease, however with relatively simple interventions this risk appears to have been mitigated.

The IQK produced results for a small sample set which are highly correlated with the results gained from both HPLC and GCMS, and was specific for the insecticide of interest. This provides a promising starting point for what could be a very useful, cost-effective method of measuring insecticide concentrations.

Chapter 4: Impact of habitat changes on tsetse populations and trypanosome prevalence in Tanga region

4.1 Introduction

From studies conducted in the Serengeti National Park, a model was developed by Lord *et al.*, (2018) which estimated tsetse abundance based on temperature, elevation, soil moisture (Landsat 8 Band 7) and vegetation (NDVI; Landsat 8 bands 4 & 5). The model was developed using parameters of environmental variables gained from the Serengeti ecosystem, in Mara region (Figure 47), however this does not represent the entire country. For models to be applicable across an entire region or country, they must be functional across a range of environments. To understand the responses of tsetse populations to environmental variables, we must undertake quantitative studies of the relationships between tsetse abundance and environmental variables in a wider range of habitats and agro-ecological settings. Combining data from all these studies will ultimately allow us to produce fine-scale predictive maps of the distribution and abundance of tsetse species across their range. Towards this goal, studies were undertaken in the Tanga region (Figure 47) of Tanzania. This region differs from the Mara region in its climate, agriculture and ecology.

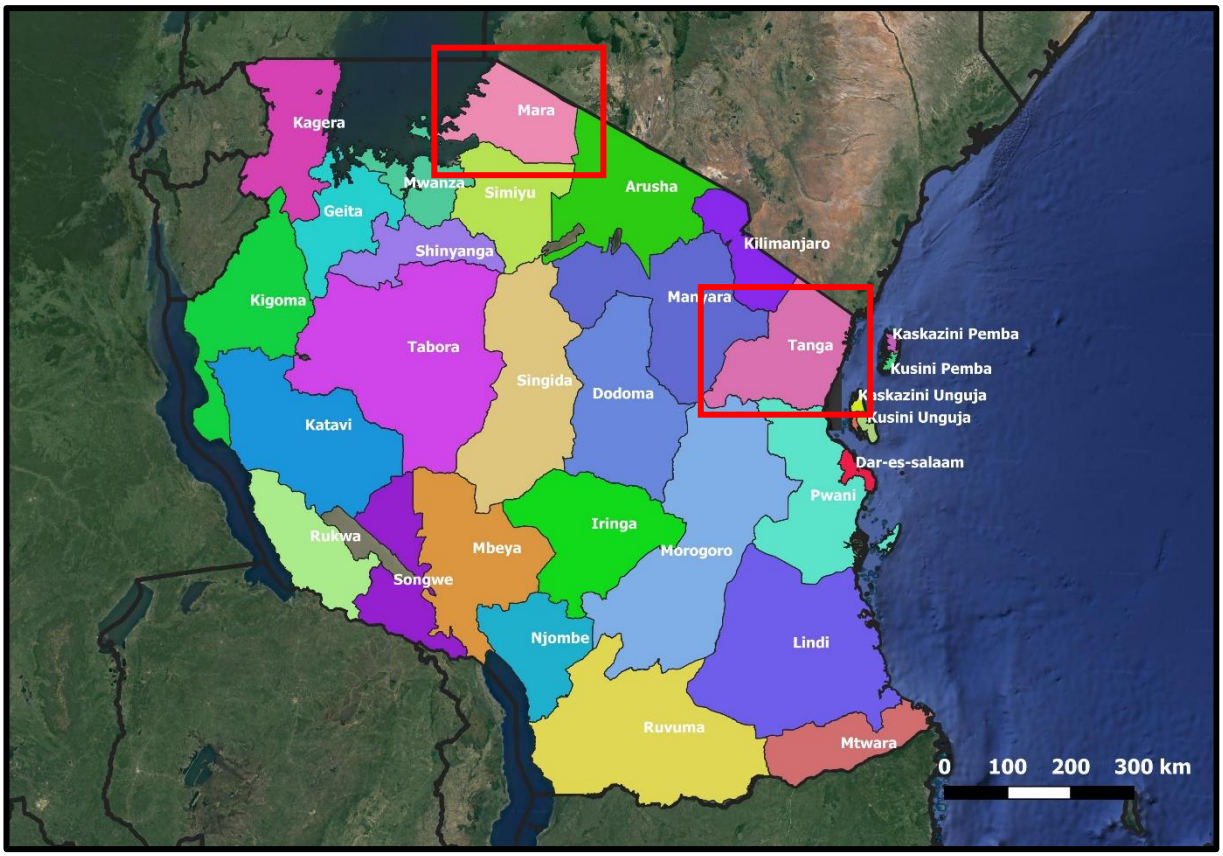

Figure 47. Regions of Tanzania with Mara and Tanga regions highlighted.

4.1.1 Climate

The average temperature in Mara region is 28.5 °C (Agriculture, 2012a) with an annual rainfall of 500 – 1,200 mm/year (Homewood *et al.*, 2002). In Tanga the annual rainfall ranges from 1,250 - 1400 mm (Lyimo *et al.*, 2004) and has a range of daily temperature between 20 – 30°C. The coast is a very humid area (Cochard, Edwards and Weber, 2015), with a mean relative humidity of 65% and range of 55 – 85% (Lyimo *et al.*, 2004).

4.1.2 Natural vegetation

In the Serengeti area, grassland plains are estimated to cover approximately 30% of the land (Reed *et al.*, 2009), with forest, woodlands and wooded grasslands also common. In a National Park in Tanga region, the Saadani National Park, the most common vegetation cover was bushland, followed by thicket-grassland mosaic, grassland then woodland. Mangrove is also a feature along much of the coastline.

Natural vegetation cover is declining in many areas across Tanzania, including Arusha (Raphael, 2018) and Kilimanjaro region (Misana, 2012). In the greater Serengeti ecosystem, the change from natural habitat to agricultural land occurred closest to the park boundary at a rate of 2.3% a year (Estes *et al.*, 2012).

There are very few studies which examine land use or vegetation change across the whole of the Tanga region so areas such as Pangani have not been examined. However, satellite images between 1980 and 2010 in Same district, which neighbours Tanga Region and contains the Pangani river basin showed vegetation cover improvement. Transition between degraded land to sparse bushland was 10% of total land change, sparse bushland to dense bushland was 6%, conversion from bushland to forest was 4.8% and conversion from dense bushland to cropland was only 4.5%. It is thought that there has been a decrease in agricultural activities as small commerce, mining and labouring on sisal plantations have become more common occupations. It is reported that young people are more likely to be migrating to Arusha or Dar es Salaam seeking work which has reduced land pressure (Ouedraogo *et al.*, 2016). If a similar trend was occurring in Tanga region this would mean that potential tsetse habitats could have persisted or even expanded in the last 30 years.

Burgess *et al.*, (1992) described the many coastal forests of Tanzania, and comments that Msumbugwe forest in Pangani District was a protected area but still at risk from people cutting for timber and fuel. The other forest areas in Pangani District however had no formal protection, nor do any forest patches in Tanga District. At the time of writing, forest damage was increasing – predominantly by human forces and Burgess *et al.*, (1992) predicted the majority of forested areas would be destroyed without increased protection. If this has continued as suggested, this would largely reduce available tsetse habitat.

4.1.3 Agriculture

In Tanga region 'other crops', including sugar cane, tea and mango, were planted on 32.9% of the available land, oranges on 19.7%, cashew on 17.9% , coconut on 11.4% and bananas on 6.32% (Agriculture, 2012b). The high percentage of fruits reflects the moist soil conditions. In Mara region, 60% of available land was planted with cereals, roots and tubers on 29%, pulses on 4% and fruits and vegetables were only grown on 1% of the available land, reflecting the drier growing conditions (Agriculture, 2012a).

In the 2007/2008 agriculture census, livestock production was also reported. Ninety-four percent of cattle in Tanga region were indigenous, 6% were 'improved' dairy breeds and 0.3% were improved beef cattle. In Mara region, the percentage of indigenous cattle was 99.5%, with 0.2% dairy breeds and 0.3% improved beef breeds. As well as the composition of breed being different between the areas, herd size was also markedly different. In Tanga region, 78% of herds had only 1 -5 cattle (Table 8) with a total of 732,130 cattle in the area.

Table 8. Tanga region agricultural census 2007/2008 for cattle herd sizes.

Herd Size	Cattle Rearing Households	%	Number of Cattle
1-5	58,457	78.29	132,435
6-10	6,195	8.30	47,121
11-15	2,662	3.57	34,934
16-20	1,936	2.59	36,458
21-30	1,681	2.25	39,636
31-40	661	0.89	23,285
41-50	545	0.73	25,329
51-60	650	0.87	36,598
61-100	925	1.24	79,923
101-150	176	0.24	23,824
>151	781	1.05	252,587
Total	74,670	100	732,130

In contrast, the 2007/2008 agriculture census for Mara region showed only 33% of herds having 1 – 5 cattle (Table 9) with a total of 1,691,118 cattle. The percentage of herds with 6 – 10 and 11 – 15 cattle was 15.5% and 10.1% higher respectively in Mara than in Tanga.

Table 9. Mara Region agriculture census 2007/2008 for cattle herd sizes.

Herd Size	Cattle Rearing Households	%	Number of Cattle
1-5	31,857	33.0	97,180
6-10	22,976	23.8	177,930
11-15	13,261	13.7	172,323
16-20	8,066	8.4	144,647
21-30	7,813	8.1	193,801
31-40	4,308	4.5	149,316
41-50	2,302	2.4	107,870
51-60	985	1.0	55,839
61-100	2,838	2.9	224,811
101-150	973	1.0	115,344
>151	1,160	1.2	252,057
Total	96,540	100	1,691,118

A survey of 25 smallholder dairy farms in Tanga region who are connected to the Tanga Dairy Trust from the wards Kange, Mafuriko and Pongwe described some farm characteristics. The smallest herds (≤5) were found closest to the town centre and these were all reared using 'zero grazing', where animals are not allowed to graze freely but, instead, they are retained in kraals where food and water is provided. As distance from the town increased, herd size also increased (up to 20) and grazing type shifted to part-time or full-time grazing. The most common types of cattle used for dairy farming were crossbreeds of Zebu and exotic breeds such as Ayrshire, Friesian or Jersey (Nkya *et al.*, 2007).

4.1.4 Tsetse and trypanosomiasis in Tanga Region

Work conducted in a variety of studies mostly during the 2000s and in the decade earlier shows that tsetse and trypanosomes were abundant and widespread across the region. Farmers in the Nkya *et al.*, (2007) study were also asked about disease prevalence and their perceived importance. Trypanosomiasis was ranked third most important after East Coast Fever and Anaplasmosis, but above worm infestations, mastitis, Foot and Mouth Disease and others. East coast fever was reported to have affected 96% of farms, and 64% by trypanosomiasis.

Knowledge of trypanosomiasis by livestock keepers in Tanga region appears to be high (Ngumbi and Silayo, 2017). A questionnaire was administered to 200 cattle owners in the districts of Korogwe and Pangani in Tanga region, as well as Mvomero in Morogoro region. In Korogwe and Pangani, 30% and 75% respectively reported presence of trypanosomiasis, and 95% and 100% reported the presence of tsetse flies and were able to correctly recognise tsetse from photos. In both Korogwe and Pangani, 95% of respondents reported using insecticide application as a tsetse control method, with 100% of those using pyrethroids. Chemotherapy was used to treat trypanosomiasis by 60% of respondents in Korogwe but by none in Pangani, however routine chemoprophylaxis was used by 45% of respondents in Pangani, and only by 5% in Korogwe.

This appears to have changed slightly from 15 years ago (Torr, Kindness and Obsomer, 2000). A study examined herds in Pangani, Muheza and Tanga Districts. There were 160 farmers who had zero-grazing herds, with an average herd size of 3, and 249 farmers with grazing herds. All farmers in the area claimed to be using trypanocides, and some were also using insecticides, but this was more common in the grazing herds. The authors also collected data on tsetse catches after setting traps in Pangani and Pongwe for 25 days. At Pangani, the average catch per day per trap was 12 flies, mostly made up of *G. pallidipes*. Very few flies were caught at Pongwe, with the average less than one fly per day per trap.

The same document also reports from Handeni district of Tanga region where 145 cattle owners were interviewed. The average herd size was 28 animals and were almost entirely zebu. Farmers ranked East Coast Fever as the most important disease in the area, followed by trypanosomiasis, then Babesiosis and Anaplasmosis. Whilst farmers had good knowledge and reported high uptake of trypanocides, Berenil in particular, very few farmers had any knowledge of how to control tsetse. Using Epsilon traps set for 16 d, they found an average daily catch per trap of 65 flies, made up mostly of *G. pallidipes*. Cattle were also sampled in this area and a trypanosome prevalence of 20% was found (Torr, Kindness and Obsomer, 2000).

In the Pangani district of Tanga Region is Mivumoni ranch, at a government-owned ranch managed by VVBDRI (Chapter 3), Lehane *et al.*, (2000) collected *G. pallidipes* from the ranch to examine groupings of mixed *Trypanosoma* infections. Flies were collected at least bi-monthly for a just over a year between February 1996 to March 1997. The average daily catch data is not provided in the paper but at least 3741 flies were caught and dissected. The overall dissection proboscis-positive rate was 17.2%. PCR-based identification of the *Trypanosoma*-infected tsetse showed that 14.8% were infected with *congolense*-type infections, 2.8% with *vivax*-type infections and 0% for *brucei*-type infections. In total, PCR identified 352 infected flies, 64% of which were single infections, 32% double infections, 4% triple infections and 1% quadruple infections. The most common mixed infection was *T. congolense* savannah with *T. congolense* Kilifi.

Alongside the Lehane *et al.*, (2000) study, Malele *et al.*, (2003) also carried out PCR-based analyses of tsetse to identify trypanosomes which were previously described as unidentified. As well as the flies collected for both studies at Mivumoni, this study also caught flies at Pangani. They caught at least 5670 *G. pallidipes* and 979 *G. brevipalpis* here, and some *G. m. morsitans* and *G. austeni* in smaller numbers.

In 1999, tsetse were still present in relatively high numbers at the Mivumoni ranch. A study conducted by Kasilagila, (2003) examined five different trap types; the biconical, NGU, epsilon, F3 and pyramidal. Traps were collected daily and fifteen replicates were done. The highest mean daily trap catch was gained using an F3 trap, which caught an average of 85.6 female *G. pallidipes* and 33.9 males. *G. brevipalpis* and *G. m. morsitans* were also caught in small numbers.

In 2005 and 2006, a forest reserve in Tanga region called Msubugwe was used as a field collection site by Adams *et al.*, (2008). This forest is next to another forest reserve called

Mgambo, and is <50 km from Saadani National Park. The total number of flies dissected in 2005 from Msubugwe was 1692, and 1331 in 2006. The majority of the flies caught were *G. pallidipes*, only a few *G. brevipalpis* were caught. The microscopy infection rate in 2005 was 6%, and 5.8% in 2006.

4.1.5 Saadani National Park

Another area of Tanga Region where tsetse were abundant is the Saadani National Park. Saadani was only established as a National Park in 2005, before which it was Mkwaja Ranch. The ranch was opened in 1954 by a private company called Amboni Ltd., and was created with the purpose of providing meat for people working on their sisal estates. At its peak, the ranch had 13,000 cattle. The vegetation on the ranch was described as a coastal forest mosaic which includes a mixture of gallery forest, areas of scrub, wooded pastures and areas with doum palm and *Acacia zanzibarica* (Fox *et al.*, 1993).

Large numbers of *G. morsitans* and *G. pallidipes* were recorded to be present, with smaller numbers of *G. brevipalpis* and *G. austeni.* No human cases of trypanosomiasis were recorded here, however flies did transmit *T. congolense*, *T. vivax* and *T. brucei* causing animal trypanosomiasis. Ranch staff were initially controlling trypanosomiasis solely with trypanocides, however in 1970 the ranch attempted to clear half the northern area of the ranch of primary bush. This very quickly became unsuccessful, as the bush re-grew rapidly due to cattle grazing on grasses which would have helped natural fires to burn, and the selectivity of their grazing left bush seeds undisturbed (Fox *et al.*, 1993).

The introduction of cattle dipping with insecticides in 1989 reduced daily catches of tsetse from ~140 to ~10 within a year and decreased the amount of trypanosomiasis but did not eliminate the disease completely. Average catches of tsetse per day remained low at approximately 5 per trap, however a study in 1999 with an improved trap method caught over 200 per day (Hargrove *et al.*, 2000). Eventually, the cost of trying to control the bush encroachment as well as the costs of treating disease left the ranch with financial deficits and in 2000 the ranch was closed (Tobler, Cochard and Edwards, 2003).

Tobler, Cochard and Edwards, (2003) analysed satellite images from the site taken in 1994 and noted radial trends in vegetation. Grass type vegetation increased with distance from paddocks, with short weed type grasses closest to cattle areas and taller grasses found further away. The ground was mostly bare in the few hundred m closest to the paddocks, then bush vegetation was found most prevalent 500 – 1000 m away before decreasing

away from the paddocks. The authors suggested that stocking rates should have been lower, and a more uniform grazing distribution used, rather than having focal points.

Treydte, Edwards and Suter, (2005) reported on ungulate populations found in the area which is now Saadani National Park. Within the park as a whole, population numbers of several ungulate species doubled between 1991 and 2001 (Figure 48), showing that as the ranch was winding down operations, natural populations of wild mammals began to benefit. An increase was not however seen in abandoned areas of the ranch where cattle had been grazed, thought to be due to poaching which occurred until 2001 when the area gained official protected status as a National Park. However, if ungulate populations have continued their trend of increasing in numbers since 2001, this would be providing many wild hosts which would not have been treated with insecticide.

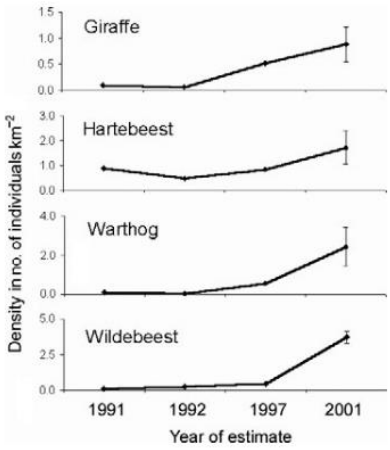

Figure 48. Density estimates from overhead flights of the four most prevalent ungulates found in the Saadani National Park area between 1991 and 2001 (unadapted from Treydte et al., (2005)).

Cochard and Edwards, (2011) investigated how woodland was developing post-ranching within the park. Woodland is associated with higher soil moisture (shown by Band 7) making it more appropriate for tsetse habitats and provide much more shade for a resting tsetse than long grasses. They studied *Acacia zanzibarica* trees within 97 study plots between 2002 and 2005 and monitored tree condition and mortality. By 2005, over a third of the study trees had died and it was concluded that the woodlands would not develop into forest. The authors describe how *A. zanibarica* would only survive when intensive grazing is ongoing around them, and unless the number of ungulates significantly increased to increase the amount of grazing, the vegetation was likely to revert to grassland. This would have negatively impacted on the availability of tsetse habitat in the area, however if

the trends shown in the (Treydte, Edwards and Suter, 2005) paper of ungulate number increasing has continued, this may have helped forests develop.

4.1.6 The predictive model

The Lord *et al.*, (2018) model which was discussed in Chapter 2.1.1 will be used to guide tsetse sampling by identifying areas of relatively high and low abundance in the Tanga region. This model was run in R using the package 'geostatsp' and the function glgm (Appendix A.3.5). It contains the environmental variables Landsat 8 Band 7, NDVI, LST elevation and this had the highest predictive power as guided by log -conditional predictive ordinates (log-CPO) scores (Appendix A.2) with the coefficients shown (Appendix A.2). This was the best model in comparison to models containing the above variables individually, with two variables or with three. A CPO score is gained by the model attempting to predict the value of points which have not been included in the fit in a step-wise process.

Catches of tsetse increased with soil moisture. Soil moisture is measured by Band 7, with a lower value indicating a higher soil moisture content. Tsetse catches also increased with NDVI values, the measure of 'greenness', with a high NDVI value indicating a more 'green' area and generally reflects the quantity of vegetation.

Other studies have analysed the effects of environmental factors on the relative abundance of tsetse. In Kenya, Williams, Brightwell and Dransfield, (1990) showed that *G. pallidipes* catches increase up to 34 °C, then decrease after 34 °C. Although air temperature is not the same as land surface temperature, the two are very closely related (Gallo *et al.*, 2011). At elevations over 1200 m in south-western Ethiopia, catches have been shown to decrease as elevation increases (Duguma *et al.*, 2015). In Zimbabwe, elevations over 1100 m were found to be unsuitable for tsetse, whereas low lying areas (lowest area ~400 m) were suitable (Matawa *et al.*, 2016). The relationships with these environmental variables may vary across the country so while models will vary between locations, the general trends are likely to be similar.

4.1.7 Chapter objectives

The climate of Tanga Region makes it particularly suitable for tsetse. The high relative humidity and rainfall provide essential habitat and conditions that protect tsetse from desiccation. Against this, the area is likely to be subject to much greater levels of habitat destruction due to higher densities of people and levels of agriculture combined with relatively small protected areas in comparison to the Serengeti National Park in Mara region.

Densities of hosts in Tanga region are also likely to be more dominated by livestock rather than wild animals and the overall densities and availability of hosts is likely to be lower; many cattle in Tanga region are maintained in a zero-grazing system which reduces their mobility and availability to tsetse. Several studies reported large catches of tsetse from the 1990s to the mid-2000s at Mivumoni, Msebugwe and Pangani. In Tanga Region it appears as though there has been a shift in recent decades from mainly trypanocide focused control to the combined use of trypanocides with pyrethroid-based insecticides. The increased and widespread use of pyrethroids may have had an impact on tsetse populations across farming areas.

Intensification of agriculture in the region may have degraded natural vegetation and contributed to a decline in tsetse abundance. This overall decline may however not apply to Saadani which changed from a cattle ranch with high densities of insecticide-treated cattle and degraded vegetation to a national park with increasing densities of natural hosts and vegetation.

Studies carried out in the Serengeti area showed the impact of vegetation on tsetse and the likely impact of use of pyrethroids on local tsetse populations. The present chapter will examine whether these same effects are apparent in Tanga region, which provides an opportunity to examine how *G. pallidipes* responds to environmental variables in an agro-ecological setting which contrasts with that of Serengeti.

Anecdotal evidence from VVBDRI staff that tsetse fly populations have declined in Tanga region may indicate that insecticide use is widespread across the country. While it was not possible to collect hair samples from local farmers, records from Mivumoni were examined to determine if insecticide may have contributed to the apparent decline in tsetse numbers there. Cattle were also sampled at Mivumoni to see if this is the case, if it has impacted on

disease transmission. Where tsetse were found, they were examined for trypanosome prevalence to gain a fuller picture of transmission dynamics in the area.

The aims of this chapter are to:

- Apply a geostatistical model of tsetse distribution, developed originally from empirical data related to the Serengeti ecosystem, to farming and wilderness areas of Tanga region to identify sites where tsetse are predicted to be abundant.
- Carry out field studies of tsetse abundance to test the predictions and develop a revised geostatistical model of tsetse abundance based on empirical data from Tanga region.
- Use satellite imagery to examine changes in land use over the past 20 years and assess whether these changes have had an impact on tsetse populations.
- Determine if insecticide use has contributed to the reported reduction of tsetse populations in at the Mivumoni Ranch in Tanga region, or contributed to a reduction of trypanosomes in cattle.

4.2 Methods

4.2.1 Satellite image processing and model testing

Satellite images were downloaded from USGS EarthExplorer (*USGS EarthExplorer*, 2019). The location was entered as path 166 and row 064 which covered most of the Tanga region down to Saadani National Park. Additional selection criteria based on Scene Cloud Cover was initially <10% and then expanded up <30%. A Landsat 8 image from January 2016 was selected as the most recent image with the least cloud cover to provide the environmental variables.

The methods used to process the Landsat 8 images are the same as in Chapter 2 (section 2.2.3). However, for the coastal area four tiles were required to cover the whole area of interest which were merged together before cropping. Each raster was then re-processed into a 500 m resolution to allow the model to produce an output for a greater area in a shorter time. This was also the resolution used by Lord *et al.*, (2018) so coefficients can be used in most comparable way. This was done by creating a grid of evenly spaced points with surrounding buffer zones and an average taken of the pixels within those buffers. The model was run in R (Appendix A.3.4) with actual data for the environmental variables for the new areas of interest in Tanga Region, with coefficients for each variable based on tsetse abundances from the Serengeti National Park to create a predictive map of tsetse at 500 m resolution. The R packages used for this were rgdal, raster, landsat and maptools.

4.2.2 Field studies of tsetse abundance

Tsetse sampling was guided by predicted maps across a range of land types, in roughly the same areas sampled in previous literature. In August and September 2017, sampling was undertaken at four main sites. Firstly, a transect was conducted extending from Mivumoni ranch in to Kibubu forest reserve. The second was conducted at Pangani river, the third at Mgambo forest reserve and the fourth inside Saadani National Park. Nzi traps were deployed and baited with acetone, 1-octen-3-ol, 4-methylphenol and 3-n-propyphenol (Torr, Mangwiro and Hall, 2011). Traps were checked and flies collected every 24 h, for three days of collection. In February and March 2018, traps were set on the same transects, again using Nzi traps for three days of 24 h collections and baited with the same attractants, however the Mgambo transect was extended and a second transect was set in Saadani National Park (Figure 49). Average catch was calculated and linear models fitted as in 2.2.1.

4.2.3 Adapting predictive model

The model developed by Lord *et al.*, (2018) was used to guide initial field sampling by identifying sites where tsetse were predicted to be relatively abundant or sparse. The predicted abundance values from the Serengeti model were compared with actual catch numbers directly and using a ranking method. A buffer was set of 350 m to get the average predicted value around each trap and these values were logged. This buffer size was the minimum size which would produce a value in R, due to averaging within 500 m pixels. The maximum logged value was set in a new column as 100%, and all the values below this were calculated as a percentage of the maximum value. The actual catches were also calculated as a percentage of the maximum actual catch. These percentages were then plotted against each other and linear models developed for the transects, individually or pooled. A Wilcoxon rank sum test was also conducted to compare actual and predicted values.

The coefficients used in the Serengeti model are based on parameters from Serengeti data, and while the general trends might be similar, the absolute values are likely to differ due to the different habitat types in Pangani and Serengeti districts. Using the actual count data and values for environmental variables from the Saadani National Park and Mgambo Forest Reserve, several new models were produced relating catch and remotely-sensed environmental variables. These models were created in R using the package 'geostatsp' and the function glgm (Appendix A.3.5). These new models were compared using the sum of their log-CPO score. The model with the highest sum of the log-CPO score was then selected to create a new predictive map for Pangani.

4.2.4 Habitat change

The distribution and abundance of tsetse in Pangani may have varied in the last 20 years due to anthropogenic changes in landcover. To analyse whether this was the case, studies were made of changes in Band 7 and NDVI from Landsat images of the region produced between 1989 and 2017; previous studies have shown that the abundance of *G. pallidipes* is strongly correlated with these bands (Matawa *et al.*, 2016). The search for suitable Landsat images was carried out using the Landsat Collection 1 Level-1 in the Landsat 4-5 TM C1 Level-1 and Landsat 8 OLI/TIRS C1 Level-1. As previously described, only images with 10-30% cloud cover were selected. There were very few suitable Landsat images for the Tanga and Saadani areas due to high levels of cloud cover associated with their coastal location. A Landsat 5 image from 1997 and a Landsat 8 image from 2016 were selected due to cloud

cover and relevance of timings. For Landsat 5 images, Band 3 and 4 are used in the NDVI calculations, rather than Band 4 and 5 for Landsat 8.

After cropping to remove NAs and clouds, Landsat 4-5 images were then converted to radiance then reflectance (Appendix A.3.1 and A.3.2) using the relevant calibration values for Lmin, Lmax, Earth-Sun distance and TM Solar Exoatmospheric Spectral Irradiances (Chander and Markham, 2003) while Landsat 8 were only converted to reflectance as described previously.

Once the images were processed, they were histogram matched to normalise the images to account for the brightness on different days. The image with the greatest contrast was used as a master image which was the image from 1997. The histmatch function (R package: landsat) was used to normalise the 2016 image to the 1997 image with smallest and largest values from both images set as the range.

After the images were normalised the values in the older raster were subtracted away from the newer image and a raster created of the change between the two images. A buffer of 250 m was set around each trap and an average value of change gained for each trap and then for each transect. A 250 m buffer was set rather a 100 m buffer (Chapter 2) to capture more of the change across a larger area as this may have had longer term impacts on tsetse populations.

4.2.5 Trypanosome prevalence in cattle

Following results gained from the tsetse collections, blood samples were then taken from cattle in one area to attempt to provide complementary evidence about the disease prevalence in the area.

4.2.5.1 FTA sample collection

To determine if trypanosomes were present in the cattle at the VVBDRI Mivumoni ranch, blood samples were collected in June 2018. The total herd size for cattle was 273 as of 06/06/18 (Table 10). The available sampling population was decided as cattle over one year of age, as studies have shown that tsetse select larger, older animals (Torr, Maudlin and G. A. Vale, 2007), which have less defensive movements. This left 181 cattle available. A required sample size of 157 animals was needed, using a test sensitivity of 0.90, a desired herd sensitivity of 0.95, and a design prevalence of 0.01. Treatment records from the previous years showed that Berenil had been used on 3 occasions, which suggested a maximum prevalence of 1%. An additional 3 cows were sampled to give a total of 160, in case of any failed tests in the laboratory.

Table 10.VVBDRI Herd demographic at time of sampling.

Animal category	Total
>1 year female	163
Female calf	51
>1 year male	18
Male calf	41
Total calves	92
Total >1 year	181
Total cattle	273

Cattle at the ranch were herded through a crush and restrained with rope to reduce the risk of injury to the cattle and cattle handlers. A qualified veterinarian then pierced the ear vein with a lancet and filled up a micro haematocrit heparinised tube (BRAND® inner diameter 1.15 mm). The blood collected in the tube was then applied to on two sets of Whatman® FTA® cards (125 µl max volume per sample area) and one glass slide.

Thin films were prepared on the glass slide and a Giemsa stain applied. Slides were then examined by trained scientists at VVBDRI, examining 100 fields per slide at 100 x magnification. One set of FTA cards was given to VVBDRI and another set returned to LSTM.

4.2.5.2 DNA Extraction from FTA

Extractions from FTA cards took place using a holepunch. Between each punch, the hole puncher was cleaned using 10% bleach, 70% ethanol and water. Holepunches of filter paper were also taken before the first sample, after each sample and after cleaning the holepunch to remove any residue and ensure the holepunch was dry. One 6mm holepunch was taken from each sample and placed into individual 1.5 mL eppendorfs. Once all the holepunches had been taken, 1 mL of distilled water was added to each tube and the tubes placed on the rotating rack to wash the holepunches for 30 min. The supernatant was then removed and another 1 mL of distilled water added before washing for another 15 min. This was repeated once more for 15 min before removing the supernatant.

A chelex solution of 0.2 g chelex, 40 µl of proteinase K and 4 mL of TAE buffer was made, and 135 µl of the solution was added into each Eppendorf. The tubes were agitated and centrifuged for 20 s. The tubes were then placed into a heat block set at 56°C for 1 h, then at 93°C for 30 min with a centrifuge step in between the two temperatures. Following

incubation they were centrifuged again and 100 µl of supernatant from each tube was transferred to individual wells in a 96 well plate before sealing with strip caps.

4.2.5.3 PCR - MPMK

A nested PCR was run with MPMK primers. For the first nest, the primers used were Tryp 3 and Tryp 4 (Adams *et al.*, 2006) and made into a mastermix with water and taq (Table 11).

Table 11. Mastermix components for Tryp 3 & 4 nest 1 to be added in the order of the table.

	X1 (µl)
Water	9.5
Tryp 3	1
Tryp 4	1
Taq	12.5

Twenty four µl of each mastermix was added to each well in the PCR plate with 1 µl of sample. In the plate a negative and a positive control was added and sealed with dome lids. Plates were then spun down in the plate centrifuge and placed on the PCR machine with the following protocol:

95˚C – 3 min
95˚C – 30 s
56˚C – 30 s } x35 cycles
72˚C – 1 min
72˚C – 5 min

For the second nest a generic forward primer was used with three specific reverse primers for *T. brucei*, *T. congolense* and *T. vivax* in a mastermix (Table 12).

Table 12. Mastermix components for MPMK nest 2 to be added in the order of the table.

	X1 (µl)
Water	9.5
MPMK2F	0.5
MPMKTBR	0.5
MPMKTC	0.5
MPMKTV	0.5
Taq	12.5

Twenty four µl of the mastermix was added to a new plate and 1 µl of product from nest 1 added to each well. The plate was then spun briefly in a centrifuge and put on a PCR machine with the following protocol:

95˚C – 3 min
95˚C – 30 s
60˚C – 30 s
72˚C – 1 min
} x35 cycles
72˚C – 5 min

4.2.5.4 Gel imaging and cleaning

To confirm the presence of a PCR product the samples were run on a gel. 80 mL of TAE buffer and 0.8 g of agarose were combined in a flask and microwaved for 1 min to dissolve the agarose. 2-3 µl of gel stain Peqgreen was then added to the flask. The contents were poured into a tray and left to set with well combs in. Once the gels were set they were placed into the tanks and the tanks filled with TAE. 2 µl of sample was then placed into each well and 2 – 5 µl of DNA ladder was placed into the first and last well of each row. Gels were run for 65 min at 105 volts. Once the gel finished running it was drained and placed into the gel imager. The programme GeneSnap (SynGene) was used to visualise the DNA in the gel to confirm if products had been produced and that the negative control had not produced a band.

Once presence of a product was confirmed through gel imaging the samples were prepared for sequencing. 0.5 µl of Exo I (Exonuclease I), and 1 µl of rSAP (Shrimp Alkaline Phosphatase) were added to 5 µl of PCR product. The mix was incubated at 37 °C for 15 min then at 80 °C for 15 min. Samples were then sent for sequencing with Source Bioscience (Rochdale, UK).

4.2.5.5 Sequencing results

Sequencing results were emailed back and opened in the programme BioEdit (Tom Hall, Ibis Therapeutics). The ends of the sequences were trimmed where a clear read could not be gained for each base, and any N values within the sequence manually confirmed. These sequences were then copied into NCBI blast and the highest result recorded with its percentage identity. Only matches over 97% were accepted as final sequencing results.

4.2.6 Trypanosome prevalence in Saadani tsetse

4.2.6.1 DNA extraction

Flies were stored in ethanol which was first evaporated using a heat block set to 56°C for several hours. They were then placed in fresh individual tubes each with a steel ball. A mix

was then made for digestion (Table 13) from the GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific), of which 200 µl was added to each fly tube.

Table 13. Quantities for digestion solution.

	X1 (µl)
Digestion Solution	180
Proteinase K	20

Tubes were then securely sealed and placed into a tissuelyser (Qiagen, Hilden, Germany). The tissuelyser was set to level three for two min. Tubes were then removed and spun down in a centrifuge at 1500 rpm for 30 s, before being placed in an incubator at 56 °C overnight with the lids sealed on.

The next morning samples are removed from the incubator and spun again in the centrifuge at 1500 rpm for 30 s. Samples were then each transferred to a new Eppendorf and 20 µl of RNase A solution added to each tube, before being vortexed and left at room temperature for ten min. 200 µl of lysis solution (from GeneJET Genomic DNA Purification Kit) was added to each sample and vortexed for 15 s. After this 400 µl of ethanol was added, vortexed again and spun briefly in the centrifuge.

The prepared lysate was then added to a GeneJET Genomic DNA Purification Column into a collection tube. The collection tubes with column in were then centrifuged for 1 min at 6000 x g RCF. The GeneJet column was then placed in a new collection tube and the solution in the old collection tube was discarded. 500 µl of wash buffer 1 was then added to each sample and centrifuged for 1 min at 8000 x g (RCF). The flow-through from this was discarded and the purification column placed back into the collection tube. 500 µl of wash buffer 2 was then added to each column and centrifuged for 3 min at 12000 x g (RCF). The GeneJet column was then transferred to a sterile 1.5 mL microcentrifuge tube. 200 µl of elution buffer was then added to the centre of the GeneJet column membrane to elute the genomic DNA. Samples were then incubated for 2 min and centrifuged for 1 min at 8000 x g (RCF). 200 µl of this extracted DNA was transferred to a new PCR plate for storage.

4.2.6.2 PCR - Tryp 1-4

Two different nested PCRs were conducted for the fly samples. The first was a nested PCR with Adams' primers (Adams *et al.*, 2008). The first nest included primers Tryp 3 and Tryp 4 and made into a mastermix with water and taq (Table 14).

Table 14. Mastermix components for Tryp 1-4 nest 1 to be added in the order of the table.

	X1 (µl)
Water	9.5
Tryp 3	1
Tryp 4	1
Taq	12.5

Twenty four µl of each mastermix was added to each well in the PCR plate with 1 µl of sample. In the plate a negative and a positive control was added and sealed with dome lids. Plates were then spun down in the plate centrifuge and placed on the PCR machine with the following protocol:

95˚C – 3 min
95˚C – 30 s
56˚C – 30 s
72˚C – 1 min
x35 cycles
72˚C – 5 min

The second nest included Tryp 1 and Tryp 2 again with water and taq (Table 15). 24 µl of mastermix was then added to each well in a new PCR plate.

Table 15. Mastermix components for Tryp 1-4 nest 2 to be added in the order of the table.

	X1 (µl)
Water	9.5
Tryp 1	1
Tryp 2	1
Taq	12.5

One µl of product from nest 1 was then taken and added to each well in the new PCR plate. Dome caps were used to seal the plate before spinning down and placed on the PCR machine and run with the following protocol:

95˚C – 3 min
95˚C – 30 s
60˚C – 30 s
72˚C – 1 min
x35 cycles
72˚C – 5 min

4.2.6.3 PCR - MPMK

To confirm results, another PCR was run with the Tryp 3 & 4 primers and the MPMK primers in another nested PCR following the same protocol used in the cattle trypanosome PCR. Gel imaging, cleaning and sequencing was also conducted as for the cattle protocol.

4.3 Results

4.3.1 Predicted model using Serengeti values for guiding sampling

The predictive map created using the Serengeti model (Figure 49) gave very high and unrealistic predictive values of abundance. The map can however be used to examine relatively high and low predictions across the selected area. The areas which were predicted to have the highest counts were around the Pangani river, at the several forest reserves and in the north of the Saadani National Park.

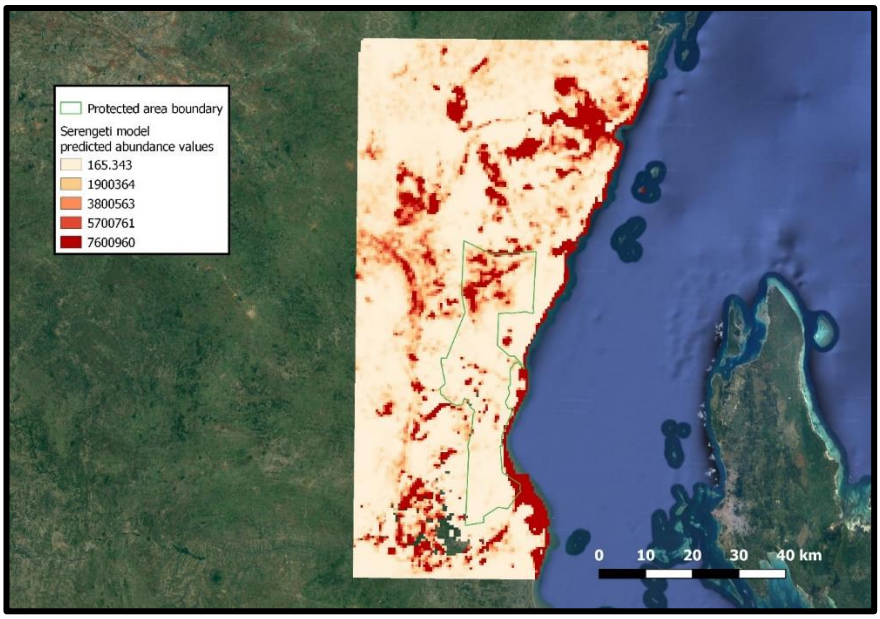

Figure 49. Map of the Tanga study area with the predictive map for G. pallidipes average daily catch created from the Serengeti model.

Based on the model predictions, a total of 62 traps were set in the study area (Figure 50). A transect of 10 traps was set starting at the Mivumoni Ranch and ending in the middle of the Kibubu Forest Reserve. Two additional traps were set slightly to the east of the transect in an area where a field assistant had reported catches of hundreds of flies per day when they worked there a decade ago. A shorter transect of eight traps was set along the north side of the Pangani river. Two transects of 12 and 14 traps were set in the north of the Saadani National Park so that they would pass through areas with a range of predicted values. At the Mgambo Forest Reserve 16 traps were placed along a road from just outside one edge of the reserve into the centre and out another side.

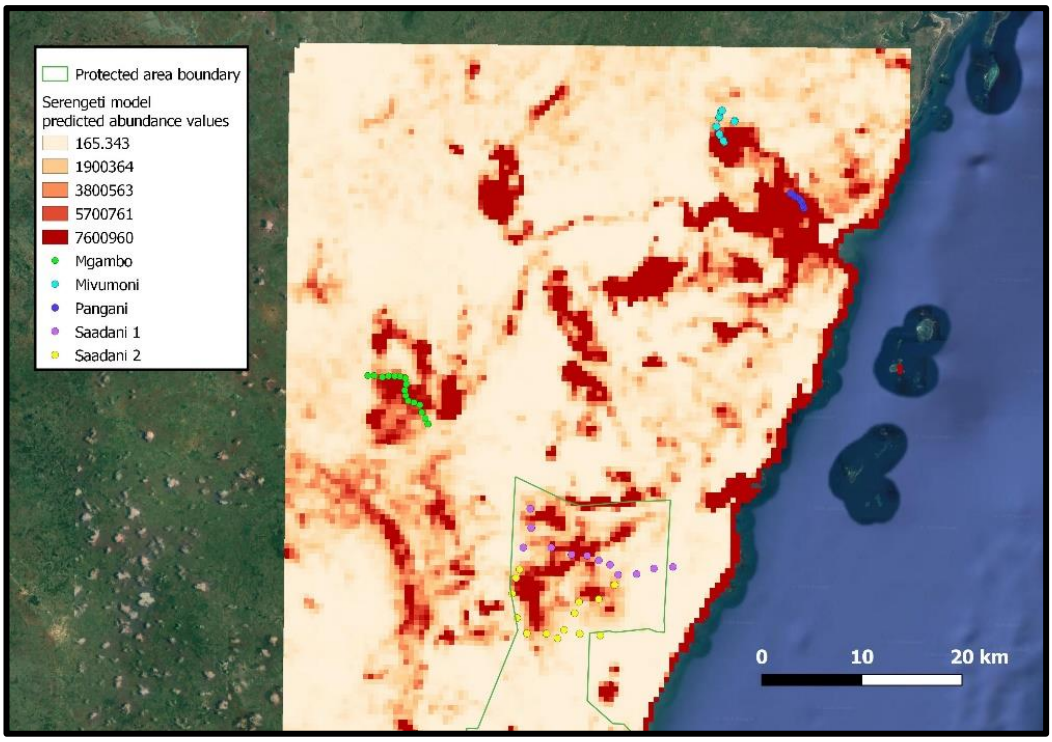

Figure 50. The location of the five transects set in the Tanga study area as guided by the Serengeti predictive model.

4.3.2 *G. pallidipes* abundance

At the Mgambo Forest Reserve the general trend showed catches lowest nearest to the edges of the forest, with the largest catches found in the centre of the forest reserve. This matches broadly with the predictions from the model (Figure 51).

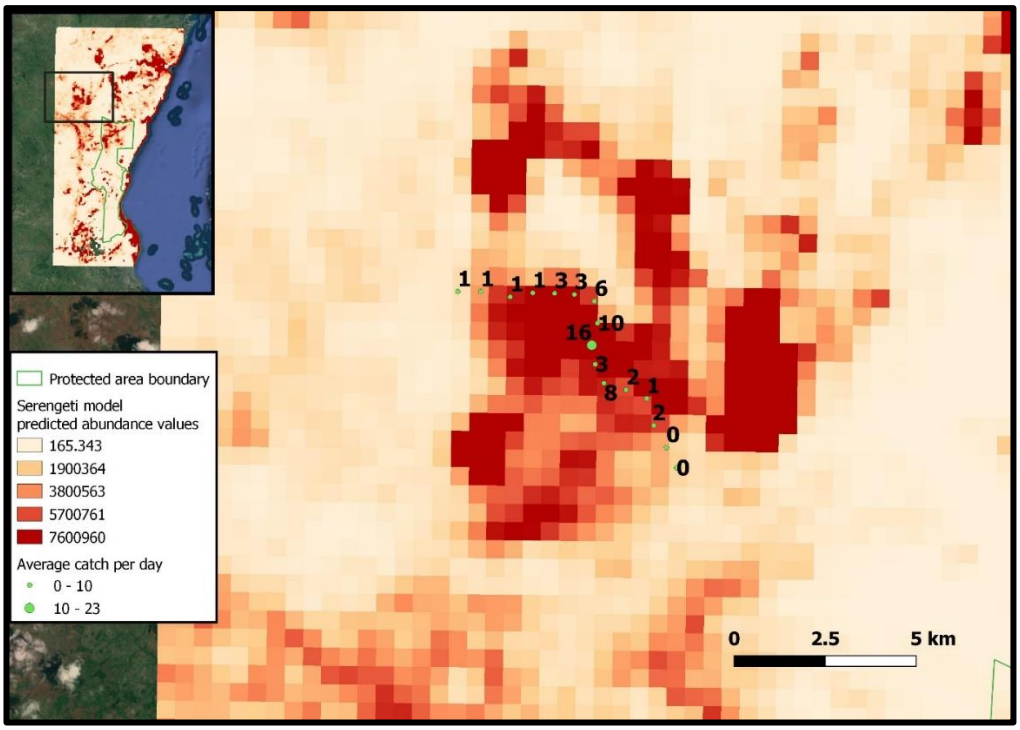

Figure 51. The Mgambo Forest Reserve transect showing the average (log detransformed) G. pallidipes daily trap catch with the Serengeti model predictive map in the background.

On the Saadani 1 transect the highest catches appear to be near to where the model suggests the highest predicted catches are found (Figure 52). The highest average *G. pallidipes* daily trap catch in the Tanga study area were found on the Saadani 1 transect (128). The average *G. pallidipes* daily trap catches on Saadani 2 transect also appeared to match the areas which were predicted to have relatively high and low counts. There does however appear to be a few exceptions to this as the southern-most trap has one of the highest actual catches (53 tsetse/day) but was not predicted to have as many tsetse compared to other areas. The trap furthest west had a moderate actual catch (21 tsetse/day), whereas the model suggested this area could provide some of the highest catches.

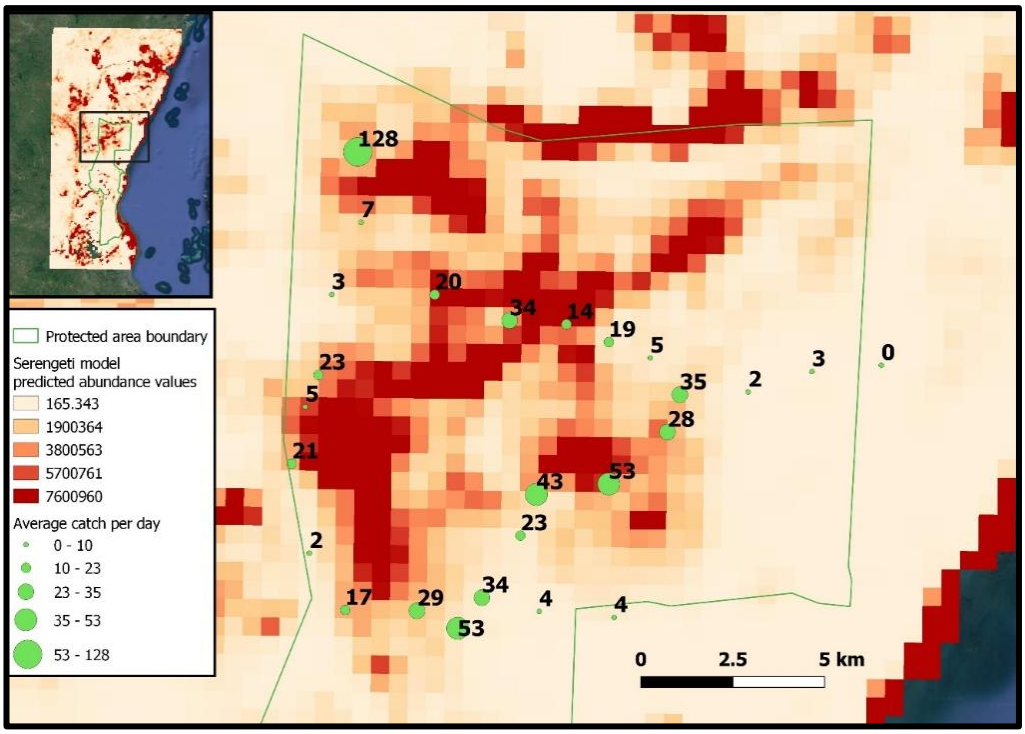

Figure 52. Map showing the Saadani 1 and Saadani 2 transect with the average (log detransformed) G. pallidipes daily trap catches with the Serengeti model predictive map in the background.

On the Mivumoni and Pangani transects, it was predicted that relatively high catches would be found along all of the Pangani transect and on half of the Mivumoni transect (*Figure 53*). However, on both the August-September 2017 trip and the February-March 2018 trip, no tsetse were caught on either transect.

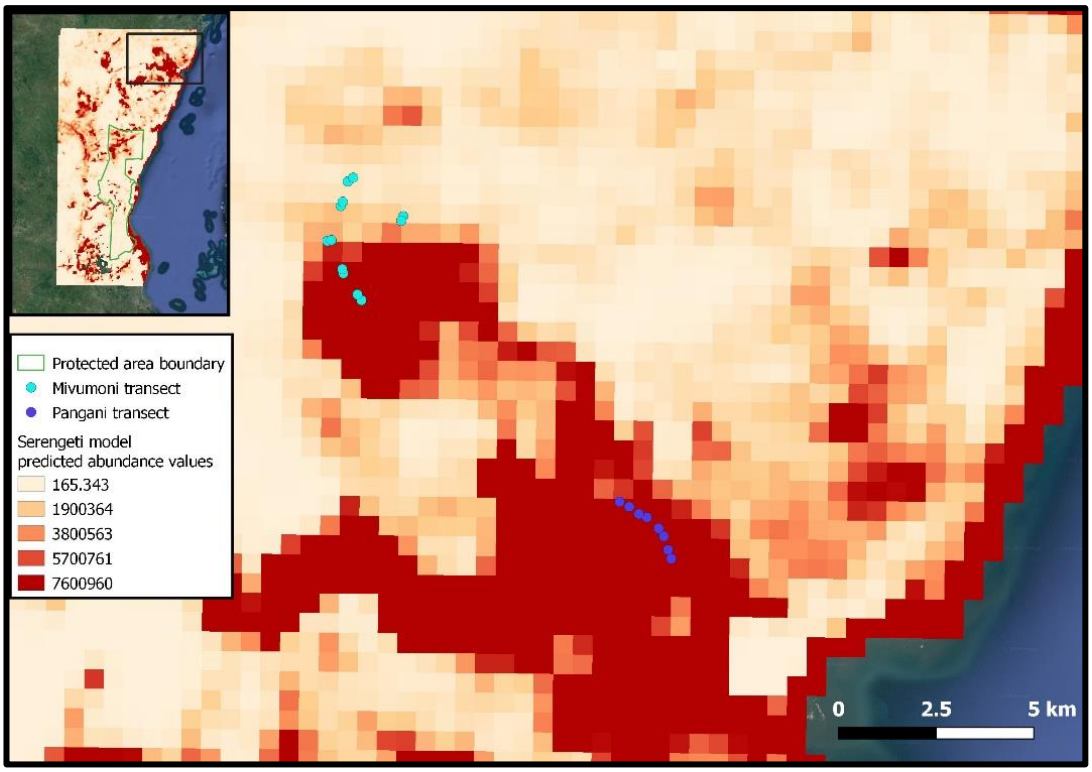

Figure 53. Transects at Mivumoni and Pangani with the Serengeti model predictive map in the background.

4.3.3 Serengeti model performance

The Serengeti model predictions were compared against catches at Mgambo, Saadani 1 and Saadani 2. Overall there was not a significant relationship between predictions and actual catches. However, when this relationship was examined by transect, two out of three showed significant correlations (Figure 54). Saadani 1 showed the most significant result (P: 0.00101, adj. R^2: 0.645), with 64% of the variability in the actual count data being explained by the Serengeti model predictions. At Mgambo this decreased to 32%, with the trend showing a similar significant relationship to Saadani 1 but over a smaller range (P: 0.0129, adj. R^2: 0.322). Saadani 2 did not show a significant relationship between the actual counts and the Serengeti model predictions. A Wilcoxon rank sum test was conducted in R which found that at Saadani 1 there was not a significant difference between the actual and predicted values (P:0.622). At Saadani 2 there was also not a significant difference (P:0.217) which appears to contradict the linear regression. At Mgambo there was a significant difference between the actual and predicted values (P:0.0155) which also appears to contradict the linear regression. At Mgambo this is likely due to the lower range of the catches compared with the predicted map, whereas at Saadani there was a greater range of catches. In Mgambo this may have been caused by a recent bush fire which could have supressed the population.

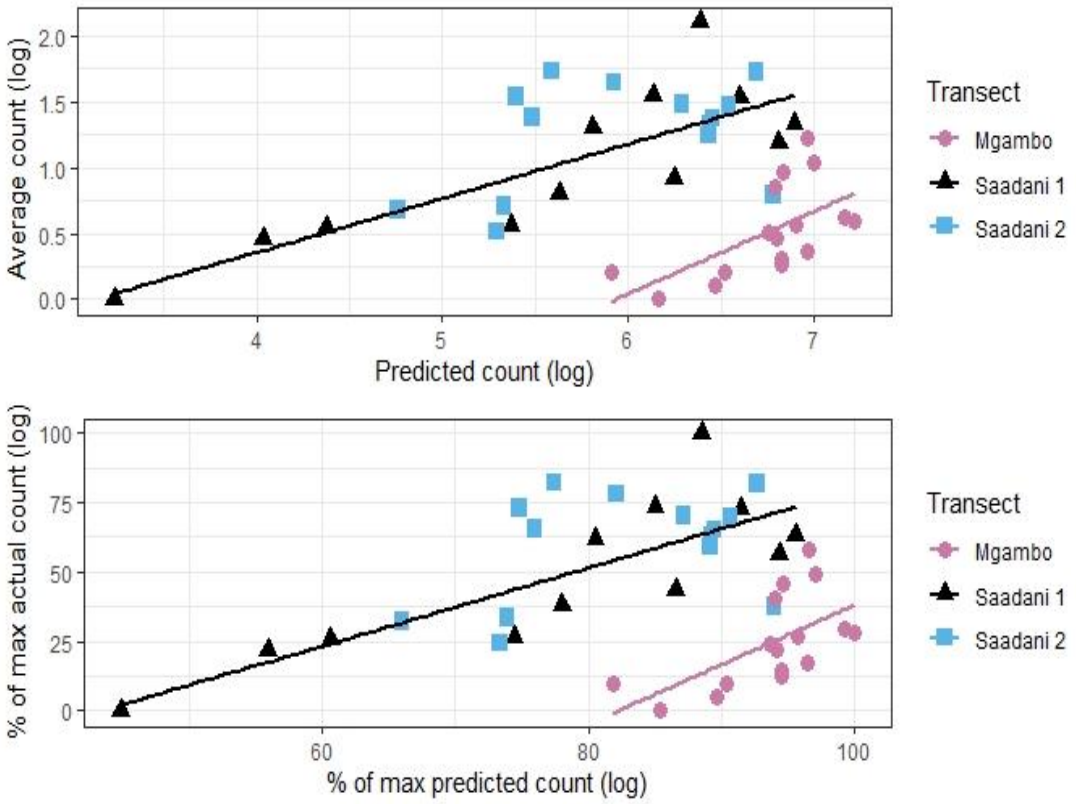

Figure 54. Predicted G. pallidipes daily trap counts from the Serengeti model compared to the actual recorded G. pallidipes daily trap counts at all transects were flies were caught, viewed by transect. In the top figure this is shown using the log of the average actual catches and the log of the predicted count. In the bottom figure this shows the values on a percentage scale, with the maximum value set as 100% and each value below that as a percentage of the maximum for both the actual counts and the predicted counts.

4.3.4 Environmental variables

The four environmental variables used in the Serengeti model – Band 7, NDVI, LST and elevation are compared with *G. pallidipes* daily trap catches from the coastal transects. The variable with the strongest effect in the Serengeti model is Band 7. When compared with the three coastal transects all together it did not show a significant relationship (Figure 55). By transect however, Saadani 1 shows a significantly negative correlated relationship (P: 0.00111, adj. R^2: 0.638) where as Band 7 values increase with decreasing soil moisture, the log average catch decreases (Figure 55). At this transect Band 7 accounted for 64% of the of variability in the data. Although Mgambo showed a similar relationship over a smaller range, this was not significant, nor was the relationship at Saadani 2.

The trends that were seen for Band 7 were also seen for NDVI. With the three coastal transects included all together in a linear model, the relationship was not significant (Figure 55). When the three transects were analysed with separate linear models, Saadani 1 had a significantly positive correlation with NDVI (P: 0.000875, adj. R^2: 0.655), with average catch increasing as NDVI increased, and was able to explain 65% of the variability in the catch

data which is 1% higher than Band 7. Again, NDVI was able to explain the least amount of variation in the catch data at Saadani 2, with Mgambo also having an insignificant relationship.

For elevation, there is a significant relationship when examined with all transects (P: 0.00258, adj. R^2: 0.186) (Figure 55). When examined by transect, Saadani 1 came closest to a significant result, however none of the transects showed a significant relationship with elevation.

LST does not have a significant relationship with trap catches when examined with all three sites together (Figure 55). However, when examined by transect, the same trend that was seen for Band 7 and NDVI is seen. Saadani 1 is the only transect which had a significant relationship (P: 0.00865, adj. R^2: 0.655) between LST and trap catches, with 66% of the variance in the data being explained by this factor. For Saadani 1 between the range of 26 °C and 32 °C, trap catches decrease as temperature increases. Saadani 2 and Mgambo did not have significant relationships with LST.

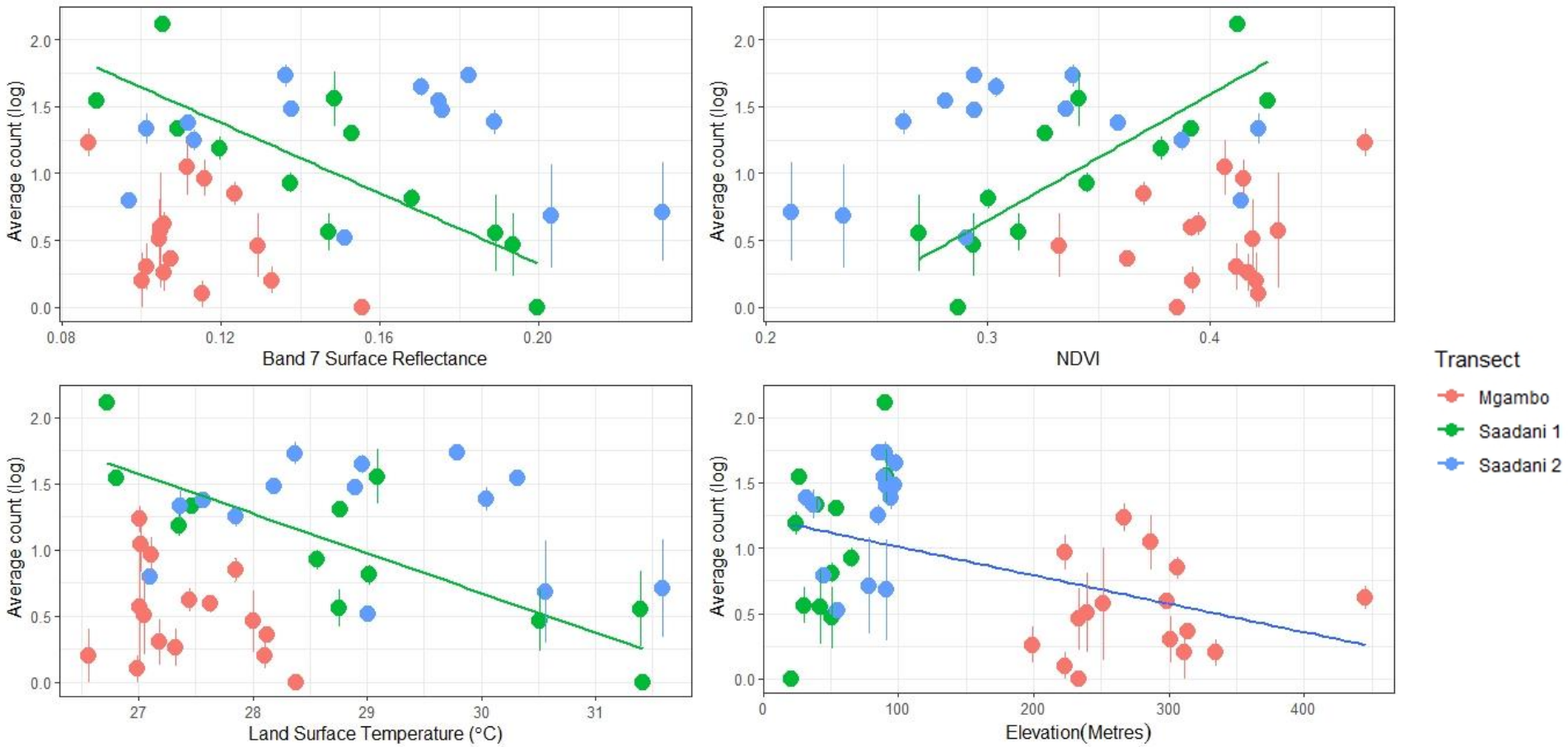

Figure 55. The relationships between Band 7 Surface Reflectance, NDVI, LST and Elevation with G. pallidipes average (log) daily trap catch. The green trendline shows significant relationships between the environmental variables with catch at the Saadani 1 transect. The blue trendline for elevation shows the relationship with all transects combined. Error bars show standard error of the mean.

4.3.5 Changes in NDVI over time

Satellite imagery was analysed to determine if the absence of tsetse at Mivumoni and Pangani was due to loss of habitat. The analysis focussed on NDVI and Landsat 8 Band 7 since these appear to have the largest impact on tsetse presence/absence. Figure 56 and Figure 57 shows changes in landcover between 1997 and 2016 with brown or green colouration indicating that the NDVI value has decreased or increased. An increase in NDVI suggests an increase in vegetation and vice versa; white indicates that there has been no change in the NDVI value for that pixel. The changes across the study area are mixed. There are large patches where the NDVI value has increased, suggesting that it is more suitable for tsetse, but there are pockets where it has decreased, including large parts of the Saadani National Park and around the Pangani river. When averages were taken from buffers around all five transects (Table 16), the total average change was positive showing a general increase in NDVI values and green-ness.

When the Mivumoni and Pagani transects are viewed in detail (Figure 56), the NDVI values appear to have decreased the most in the riverine vegetation surrounding the Pangani river. At the Mivumoni transect, the NDVI values near the southern part of the transect in the Kibubu forest appear to have decreased slightly but these changes appear faint and large parts have remained the same. In the area outside the forest in the northern part of the transect the NDVI values have increased. The total average change at Mivumoni was positive, but the total average change at Pangani was negative (Table 16), and when the two are combined it gave a negative change.

Table 16. Average change in NDVI values taken from 250 m buffers around the traps in each transect.

Area	Change in NDVI value
All	0.0425
Mivumoni	0.0361
Pangani	-0.0914
Mivumoni & Pangani	-0.0149
Mgambo	0.151
Saadani	0.0197
Mgambo & Saadani	0.0700

For Mgambo and both Saadani transects combined, NDVI has increased with values at Mgambo showing the largest improvement. Mgambo is seen on the edge of a large area

where NDVI has increased the most across the study site (*Figure 57*). In the northern part of the Saadani National Park around the Saadani 1 transect, NDVI has generally increased, however around Saadani 2 it generally looks to have decreased. Both Saadani transects combined give a positive change (Table 16).

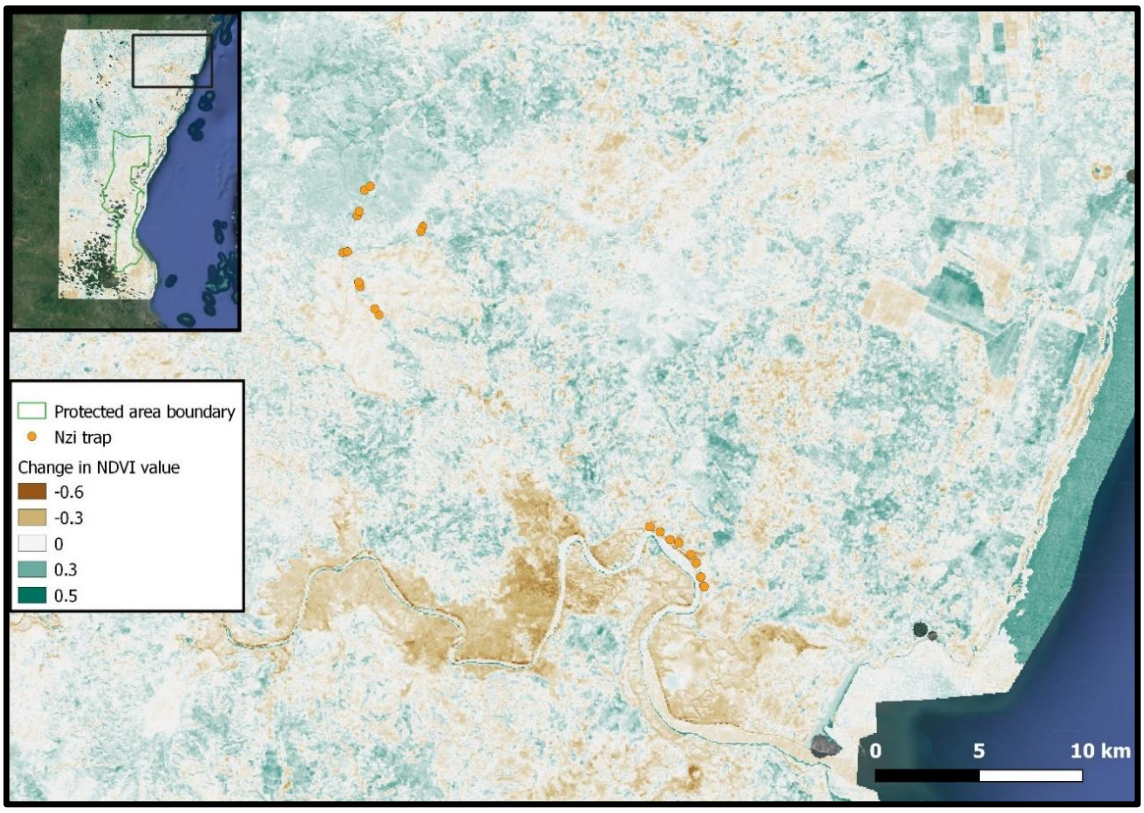

Figure 56. Map showing the change in NDVI values over time between Jan 1997 and Jan 2016 in the area around the Mivumoni and Pangani transects.

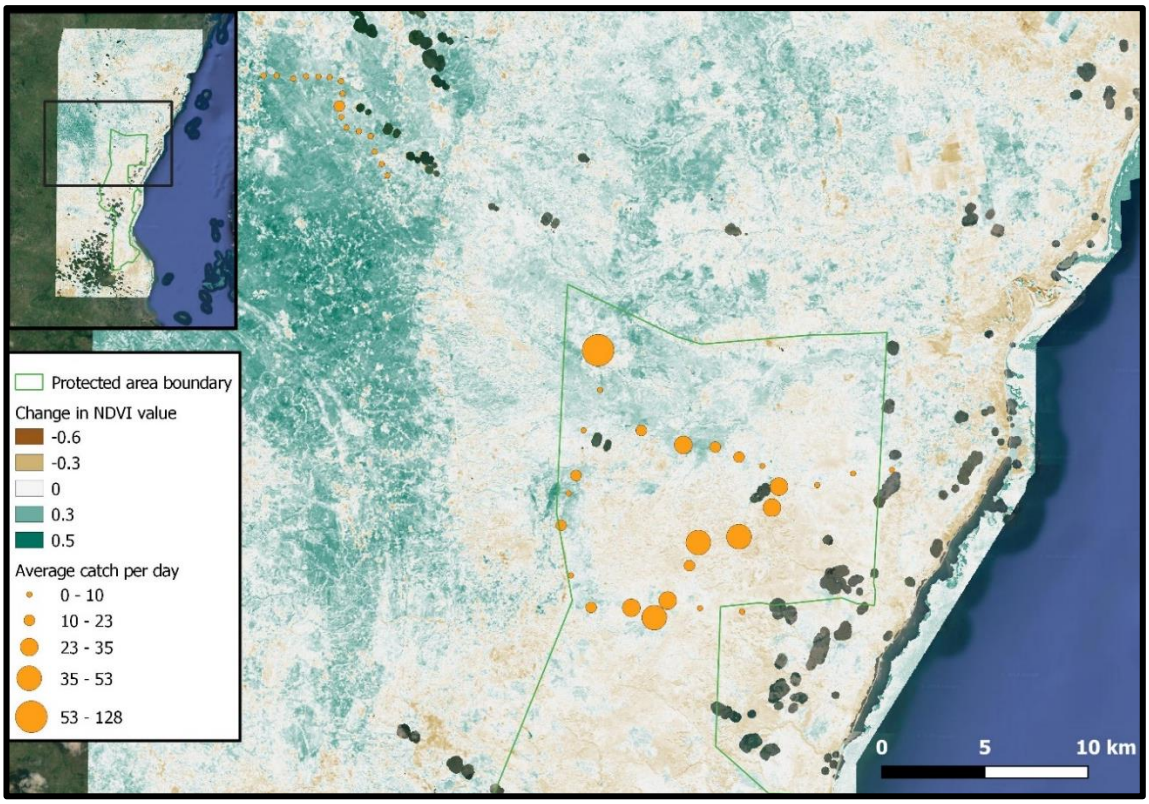

Figure 57. Map showing the change in NDVI values over time between Jan 1997 and Jan 2016 in the area around the Saadani and Mgambo transects.

4.3.6 Change in Band 7 over time

Band 7 was an environmental variable which was significant at Saadani 1. As a lower Band 7 value means a wetter site, if the change is negative this means the site has become more suitable for tsetse. A map of Band 7 value change was created for the whole study area (Figure 58 and Figure 59). Green and brown colouration indicates an area has become drier or wetter reseptively. As for NDVI, it appears that changes have been mixed across the area, however most of the Saadani National Park appears to be green. The total average change across the five transects is negative, indicating that it has come more suitable for tsetse since 1997 (Table 17).

At the Mivumoni transect, a similar pattern to the NDVI results was seen (Figure 58). The northern part of the transect is in an area where Band 7 values have decreased, and the southern part of the transect looks like it has changed very little, there the overall change for the transect was negative indicating that it had become more suitable for tsetse. Pangani was again the transect which showed an opposite effect, as the overall change was positive. For both Mivumoni and Pangani combined, the average change was negative.

When Mgambo and both Saadani transects are considered together, an overall negative change in Band 7 was recorded (Figure 59, Table 17). Mgambo had a more negative average change in Band 7 although only slightly higher than Saadani. In contrast to the NDVI change, the majority of the Saadani 1 and 2 transect areas show a negative change.

Table 17. Average change in Band 7 values taken from 250 m buffers around the traps in each transect.

Area	Change in Band 7 value
All	-0.0391
Mivumoni	-0.0530
Pangani	0.0163
Mivumoni & Pangani	-0.0253
Mgambo	-0.0471
Saadani	-0.0447
Mgambo & Saadani	-0.0456

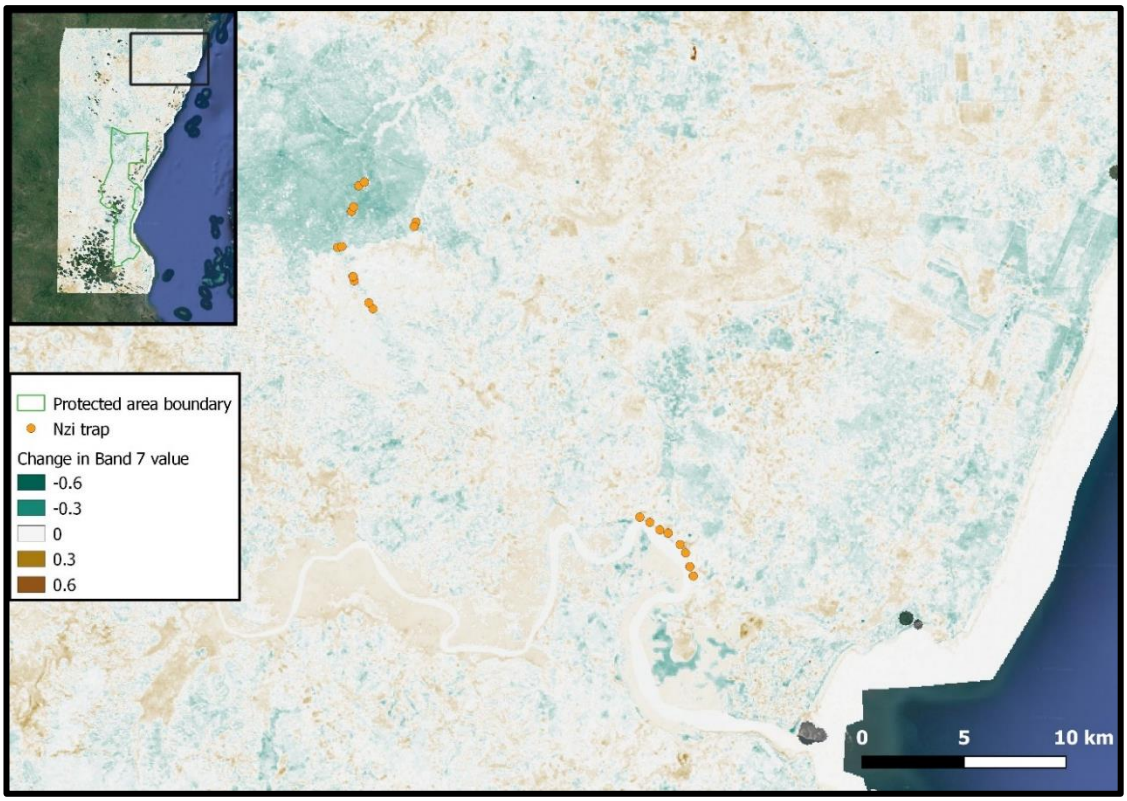

Figure 58. Map showing the change in Band 7 values over time between Jan 1997 and Jan 2016 in the area around the Mivumoni and Pangani transects.

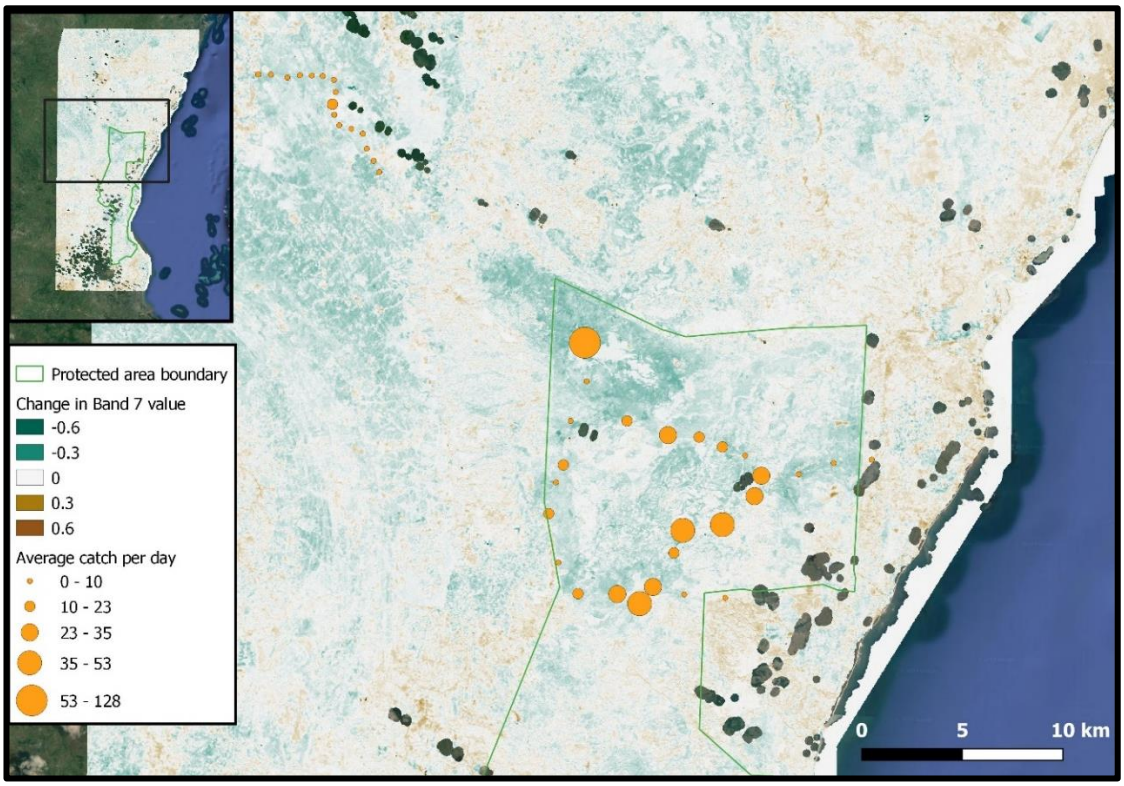

Figure 59. Map showing the change in Band 7 values over time between Jan 1997 and Jan 2016 in the area around the Saadani and Mgambo transects.

4.3.7 Adapting the predictive model

The lowest log-CPO score for models with only one environmental variable was for the model containing only Band 7 which was the minimally adequate model (Table 18). For models containing two environmental variables the lowest log-CPO score was for the model containing Band 7 and elevation, however this did not have a better predictive power than the model containing Band 7. Further combinations with more variables were tested for completeness and none scored better than the Band 7 only model. A null model was also tested to ensure that Band 7 had a better predictive power than a model containing no variables which it did. The model with the worst predictive power was one containing all the variables, which performed worse than the null model. This is in contrast to the Lord *et al.*, (2018) model, which contained all environmental variables and had the highest log-CPO score (Table 18). The posterior estimates of the Band 7 coastal model coefficients are shown in (Table 19), and has a range of 4,703 m after which distance the correlation becomes smaller than 2 standard deviations.

Table 18. Negative of the sum of the log conditional predictive ordinates (log-CPO score) for various linear models using catch data and environmental variables from the Mgambo, Saadani 1 and 2 transects.

Variable in model	Log-CPO score
Null	19.701
B7	17.0115
NDVI	19.107
Elevation	20.712
Temp	18.410
B7, temp	18.263
B7, NDVI	18.344
B7, elev	17.946
b7, elev, ndvi	19.221
b7, elev, temp	19.523
b7, elev, temp, ndvi	21.190

Table 19. Posterior estimates of model coefficients for the Band 7 coastal model which had the best log-CPO score.

Band 7 model coefficient	Mean
Band 7	-5.393
Intercept	1.685
Range (m)	4,703

Table 20. Posterior estimates of model coefficients for the coastal model which contained all environmental variables.

Model all variables coefficients	Mean
Band 7	-5.804
Elevation	-0.00183
LST	-0.0888
NDVI	-2.0201
Intercept	5.214
Range (m)	3,840

4.3.7.1 Predictive map from Band 7 model

From the Band 7 coastal model which had the highest log-CPO score, a predictive map was produced (Figure 60). To compare the new predictive map with the map produced from Serengeti data, absolute numbers predicted are less important than the relative differences in catch. In the Band 7 coastal model map the contrast is less obvious between areas where many flies are predicted to be caught, compared to areas where less are predicted to be caught than in the Serengeti model map. However, the two maps appear to highlight many of the same areas within the study area.

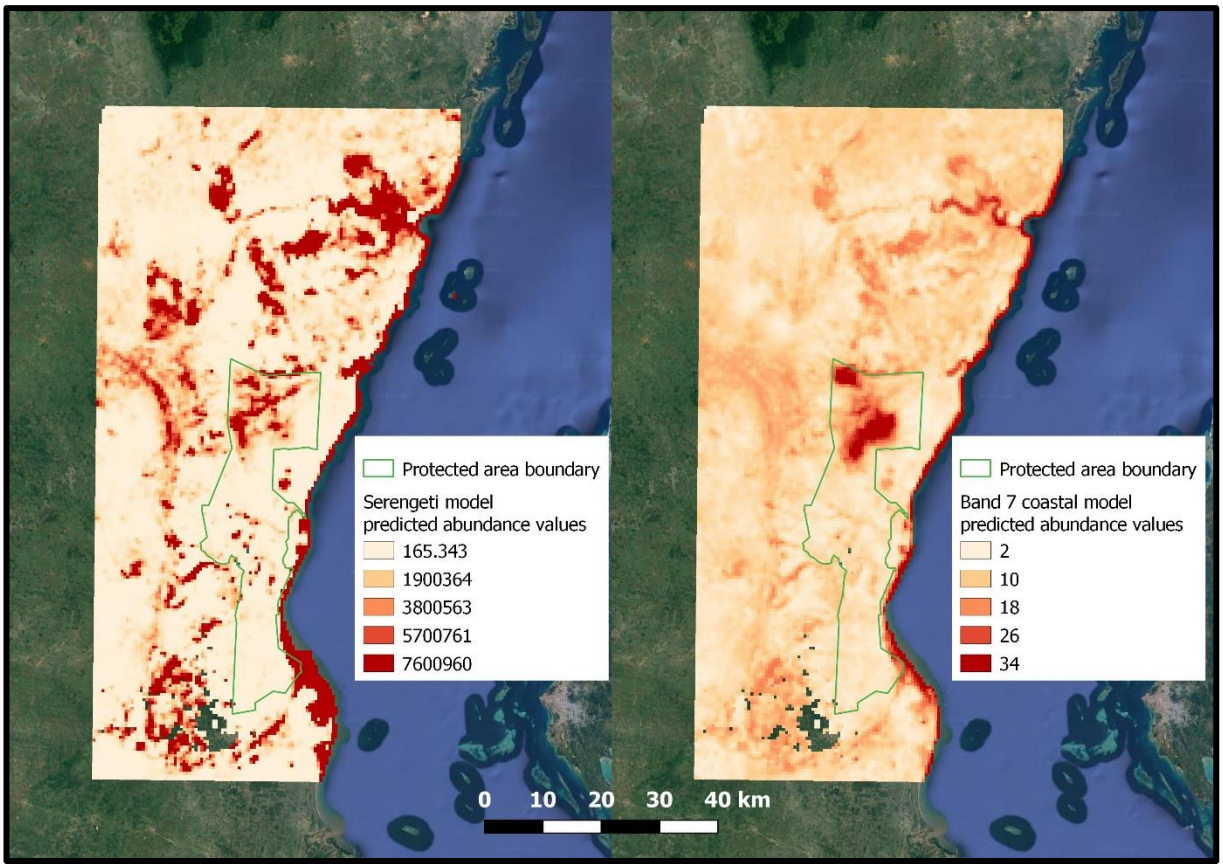

Figure 60. Map to show comparison between the predictive map developed by (Lord et al., 2018) using coefficients from the Serengeti, alongside the predictive map produced from the Band 7 coastal model. The maximum value in the Band 7 coastal predictive map was 72, however the scale has been altered to highlight areas other than Saadani where relatively high – moderate numbers of tsetse are predicted to be caught.

4.3.8 VVBDRI ranch records

On a visit to the VVBDRI ranch in 2018, insecticide records were collected from the ranch manager (Table 21). Records were not extensive, but each time a new insecticide was used this was recorded, and it is assumed that this insecticide was used until a different insecticide is named. Stored records could only be gathered from 2006 to present, but they show that the type of insecticide used was often effective against both ticks and tsetse.

Table 21. Records collected at the VVBDRI ranch. The table shows the date a change of insecticide was recorded. '<' denotes that the treatment had been used at some point from the previous date, up to this time point.

Date	Insecticide name	Insecticide group	Effective against
09/08/2006	Dominex	Alpha-cypermethrin	Tsetse and ticks
31/12/2011	Paratik	Alphamethrin	Tsetse and ticks
24/04/2015	Paratik	Alphamethrin	Tsetse and ticks
<20/01/2018	Tixfix	Amitraz	Ticks
20/01/2018	Vectorclor	Cypermethrin	Tsetse and ticks

Records were also collected from the VVBDRI veterinarian of the cattle health for the last four years. The herd size during these years was approximately 300 - 350 cattle. The number of cattle sampled between 2014 – 2016 was ~ 30 – 55 cattle once or twice a year. Quarterly sampling began at the end of 2016 but only 2 – 5 cattle were sampled each time. Blood samples were taken and examined for pathogens, and stained with Giemsa to look for trypanosomes. *Theileria parva*, the causative agent of East Coast Fever and *Anaplasma marginale*, the causative agent of anaplasmosis were detected but no trypanosomes were detected in these four years of records.

4.3.9 Trypanosome prevalence in Mivumoni cattle herd

Of the microscope slides stained with blood from the Mivumoni herd which were examined by scientists at VVBDRI, none were found to contain trypanosomes. After conducting PCR on the FTA cards with the generic primer sets, 15% (n=24) of samples were positive for *Trypanosoma* sp.. Identification of the trypanosomes from the gel photos estimated *T. brucei* to be the most prevalent species in 5% of samples, followed by *T. congolense* kilifi (3.75%) and *T. congolense* savannah (2.5%). Five samples (3.13%) showed bands corresponding to other trypanosomes such as *T. godfreyi, T. simiae, T. simiae Tsavo, T. grayi* and *T. theileri.* These 24 positive samples were then run with species specific primers for *T. brucei, T. congolense* and *T. vivax*. Only 8 of the samples (5%) were amplified and these were sent for sequencing. The cut-off for certainty of sequencing results is normally set at 97%. Only four samples returned a certainty over 97%.

Final prevalence values based on sequencing report the total prevalence of pathogenic trypanosomes is 1.25%. One sample was confirmed as *T. brucei* (0.63%), one was *T. vivax* (0.63%), and two as *T. godfreyi* (1.25%).

4.3.10 Trypanosomes in tsetse

From PCR, *Trypanosoma* sp. were found in a total of 17% out of 421 samples (n=73). PCR products were seen in 18% (n=38) of 210 male flies and 17% (n=35) of 211 female flies. *T. brucei* was in 8% (n=17) of female flies, and 6% of male flies (n=13). *T. congolense* was detected in 9% (n=18) of female flies, and in 11% (n=23) of male flies. T. vivax was found in 1% (n=2) of male flies, but not in any females.

All of the samples which reported as positive from PCR were sent for sequencing. Only 22 of these had a certainty of 97% or above. The prevalence of pathogenic trypanosomes from all tsetse sampled is 4%. *T. brucei* was found in 2% (n=9) of all flies, *T. congolense* in 1.4% (n=6) of all flies and *T. vivax* in 0.48% (n=2) of all flies. *T. godfreyi* was found in 1.2% (n=5) of all flies.

4.4 Discussion

4.4.1 Abundance and environmental variables

Comparison of satellite data from the coastal district of Pangani with the inland district of Serengeti highlighted coastal areas which were densely vegetated and had values for Band 7 and NDVI indicating their suitability as tsetse habitat. Entomological surveys were conducted in areas where: (i) the abundance of tsetse was predicted to be relatively high (ii) surveys were logistically feasible and (iii) historical data indicated that tsetse had been present. Traps were placed closer together in Pangani than at Mivumoni as the mangrove habitat to the East of the Pangani transect reduced the available area to place traps. No tsetse were caught from these transects suggesting that tsetse abundance has declined markedly over the last 20 years. The lack of tsetse caught in any traps from these transects does not however mean that tsetse are not present. Stationary odour-baited traps only catch an estimated 1% of all *G. pallidipes* which approach a trap (Barclay and Hargrove, 2005), and it is possible that a very small population may be sustained the area.

The low incidence of trypanosomiasis in cattle from Mivumoni, and the absence of routine use of trypanocides also suggests a very low density of tsetse. However, the finding of trypanosomes in a few individual cattle suggests that tsetse are present. The historical presence of tsetse in habitats known to have tsetse historically (Mivumoni, areas close to the Panagani river) suggests that the Serengeti-based model was a useful means of identifying potential habitat. However, as the model only contains environmental variables, it would be unable to account for any other changes in the area such as vector control or declines in host availability through say poaching of wild hosts or destocking.

The northern part of the Saadani National Park was also selected for entomological surveys because the Serengeti-based model indicated it was also highly suitable for tsetse with strong evidence that tsetse were abundant there in the 1990s and, as a recently-designated national park, it seemed likely that vector control had not been carried out recently and natural hosts would be abundant. In contrast to the Pangani and Mivumoni transects it is the only trapping site in a National Park, which provides a greater level of protection from grazing or deforestation for wood supplies than in a Forest Reserve. When the Mkwaja ranch was near to closing, tsetse numbers had been reduced to low levels (<10 tsetse/day) due to intensive vector control through the use of ITC. During the present study, catches have increased and the highest catch of 128 (log detransformed daily average trap catch) from the whole study area was found here.

For the Saadani 1 transect; Band 7, NDVI and temperature were all significantly correlated with the average daily catches. This may explain why the relative Serengeti model predictions correlate so well with the actual catches (**Error! Reference source not found.**). The Saadani 2 transect was however not significantly related to any of the environmental variables and did not fit well with the Serengeti-based predictions. The apparent absence of statistically significant correlations between environmental predictors and catch may be subject to additional factors. The first could be due to weather. During the week of the Saadani 1 transect, the weather was hot and dry, and reflected a typical dry season day which would have been reflected in the selected in the Landsat 8 images used for analysed from January 2016. The Saadani 2 transect was conducted the following week and during this week there had been some rainy spells and was more overcast. This could explain why there are higher catches in traps which according to the Landsat 8 image, where in areas with lower NDVI values and higher Band 7 values. The Landsat 8 image from a dry season may no longer reflect the actual conditions, as vegetation may have begun to grow quickly and flies may have dispersed further with fewer hours of direct sunlight due to clouds. Another explanation could have been due to the presence of people and/or cattle. During the field sampling the presence of houses along the west side of the Saadani 2 transect was noted, and a larger village is present in the South-Western corner of the transect.

Mgambo was also selected for an entomological survey because it is a Forest Reserve where the model had predicted relatively high catches. It was further inland than the other sites and the field research team reported that flies had been caught there in recent years. Tsetse were caught but the range of daily average catches and maximum catch was much lower at Mgambo (average: 4, range: 0-16) than Saadani 1(average: 22, range: 0-128) or Saadani 2 (average: 24, range: 2 – 53). The lower catches may reflect an overall lower density of tsetse related to fewer hosts and/or overall habitat size; Forest Reserves are less protected from grazing than National Parks. During the field studies, few cattle were seen and there was no evidence of cattle having been in the area. Despite the fact that catches at Mgambo did not show any significant relationships with any of the environmental variables, the overall trends were similar to Saadani 1, albeit over a reduced scale in terms of both the variable and the total catch size. The actual catches had a significant correlation with the predicted catches from the Serengeti model, although had an adjusted R^2 value that was roughly half (adj. R^2: 0.32) that found for Saadani 1 (adj. R^2: 0.64).

When the three transects were pooled to examine each environmental variable, only elevation showed a statistically significant relationship. This significant result is likely an

artefact of site location as Mgambo is higher but had lower maximum catches. When the transects were examined individually for elevation, the slope of the line went from a negative one when grouped, to three separate positive slopes. The effect of site may also have caused the other three variables to show non-significant relationships. If Saadani 1 is taken as an example due to its large range of catch sizes and large range of environmental variables, the environmental variables show results most similar to those found in other literature. A strong correlation between *G. pallidipes* and Band 7 has been reported on several occasions, and due to the relatedness of Band 7 and NDVI it is unsurprising that NDVI was also significant.

The Serengeti-based model predicts very high numbers (>1000 tsetse/day) of tsetse in some parts of the coastal study site. The main difference between the Serengeti site and the coastal site is that overall the elevation along the coast is much lower. The range of elevation at the Serengeti is between 1200 m – 1500 m whereas at the coastal site it is between 0 – 500 m. As the coefficient for elevation in the Serengeti model is -0.003 it has a negative relationship which means this variable will predict higher numbers at lower elevations.

4.4.2 Changes in environmental variables over time

NDVI and B7 were both selected as variables to examine for change over time due to their relationship with catches of *G. pallidipes* and significance with the Saadani 1 transect. Whilst temperature was also significant, a check of yearly climate data for Tanga showed only slight temperature fluctuations over the last twenty years (Figure 61).

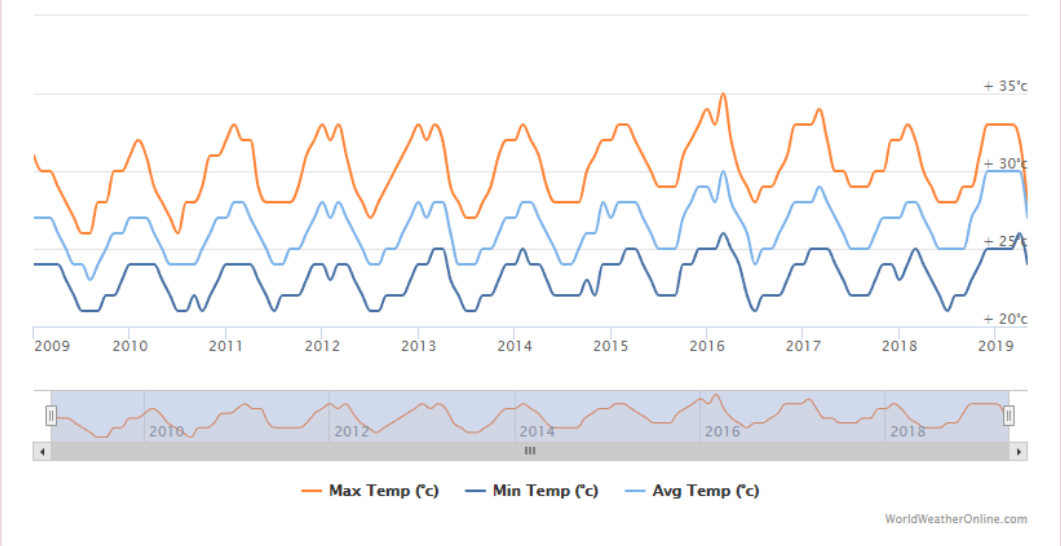

Figure 61. Average, min and max temperatures for Tanga for the past 10 years. Adapted from (*Tanga, Tanzania Weather Averages | Monthly Average High and Low Temperature | World Weather Online*).

From a search of available Landsat images between 1990, ten years before Mkwaja ranch was closed, to 2018, the two timepoints to examine change between were selected as 1997 and 2016. There were only approximately ten images which were sufficiently cloud-free for the areas of interest and 1997 was selected as this was only a few years before the ranch shut, so any habitat degradation due to the ranching operation would have been more pronounced than any images taken after it had closed. This was also around the time that Lehane *et al.*, (2000) was catching thousands of flies from Mivumoni. Ideally a selection of images would have been available around a short time period to take an average value for that year but unfortunately due to the cloud cover this was not possible. However, in the Band 7 change image, the profile of Saadani National Park stands out as an area which has obviously changed. This highlights that the changes between 1997 and 2016 are ones which are likely to have occurred over a long time period, rather than just a seasonal difference. If it had been seasonal it would be expected that positive changes could have occurred at the same rate everywhere, but the additional growth in Saadani can be seen clearly.

Averaging over the Mivumoni transect shows an improvement in terms of tsetse habitat suitability for both NDVI and Band 7. This would suggest that the decrease in tsetse catches cannot be attributed to changes in habitat. It is much more likely that increase in the use of vector control both at the VVBDRI ranch and in the surrounding villages have contributed to this, alongside a reduction in wild host densities which have been reported by VVBDRI staff (I. Malele, pers. comm.).

At Pangani, a reduction in habitat suitability was shown for both NDVI and Band 7. This could therefore be attributed to the decrease in tsetse catches in this area, as well as vector control. There are fewer reports of past or current wild host densities in this area as there are no protected areas here, however from visual inspections, several small herds of cattle were witnessed along this transect.

The most complicated situation may be at Saadani. Catches there have increased since the ranch closed and the area became a national park. Before the ranch shut there was intense periods of control through ITC which will no longer be ongoing inside the borders of the park, and there will also have been a likely increase in wild hosts. Overall it appears as though the change in values for NDVI and Band 7 would have improved the extent of habitat suitable for tsetse. In some areas where the Saadani 2 transect was carried out, the values for Band 7 have increased but NDVI has decreased. The two variables are normally positively correlated, but it could be that the type of vegetation in these areas have

different reflectance values. Tobler, Cochard and Edwards, (2003) classified vegetation types in the park using an image from 1994, which clearly shows a divide in vegetation type which roughly follows the divide in the increase and decrease of NDVI. The top left portion in 1994 comprised predominantly types of grassland, so if the ranch had shut and this grassland was no longer being grazed it would be expected to have grown and NDVI values and Band 7 values would increase as evidenced between 1997 and 2016. The bottom right portion where Saadani 2 passes through was predominantly bushland. If the bushland had continued to grow in these areas, it could be a 'less green' type of vegetation which is still replacing grasslands so the NDVI value goes down, however the fact there is more taller vegetation would probably provide more shade and moisture, so Band 7 will still go up (Browning *et al.*, 2018).

4.4.3 New model

Unlike in the Lord *et al.*, (2018) model, the best model as judged by a log-CPO score was not one which contained all the environmental variables but instead only Landsat 8 Band 7. This may be because many of the trap sites in Saadani may have confounding effects which have obscured more subtle environmental changes, but Landsat 8 Band 7 may have a greater effect on abundance. The Serengeti model was created using data from traps which were at least 10 km away from the border inside the Serengeti National Park. In the Northern section of the Saadani National Park it is only ~ 15 km at its widest. Consequently, the survey traps, even those in the centre of the park, may be affected by factors outside the park such as control operations and/or habitat degradation.

There may be evidence from other disease systems which have hinted at the impact of insecticide use in Tanga Region. Meyrowitsch *et al.*, (2011) reports that the burden of malaria in the region has significantly reduced, even though there has been no formal vector control programmes. Between 1998 and 2001, *Anopheles gambiae* and *Anopheles funestus* caught per trap decreased by 76.8% and 55.3% respectively. Between 2003 – 2009, the same species decreased again by 99.7% and 99.8%. The authors concluded that levels of monthly rainfall were likely to be linked to the decrease in mosquito numbers, however also commented that the situation was likely to be more complex, with many possible factors contributing, including the use of 'insecticide-like compounds not directly applied for targeting malaria vectors'.

4.4.4 VVBDRI ranch records

At the VVBDRI ranch, the majority of the insecticide formulations used by the ranch in the last few years are capable of killing tsetse, which provides support to the hypothesis that the apparent lack of tsetse seen in the area could be due to insecticide use.

4.4.5 Trypanosomes in cattle and tsetse

After sampling twice for tsetse on the Mivumoni transect and finding no flies, and from speaking with staff who are living and working on the ranch about the presence of tsetse it appeared as though tsetse had been eliminated from the area. The staff were no longer concerned about trypanosomiasis as they had been in previous years, and did not treat the cattle herd with prophylactic drugs. The initial microscopy also accorded with this narrative when no trypanosomes were seen on any of the slides. The high prevalence of bands (24/160=15%) which appeared during PCR was therefore unexpected. When the initial positive 24 samples were re-analysed using species-specific primers only eight (5%) of these were amplified. When re-attempting to run these 24 samples again, some of the same samples re-amplified, but others did not, and others which had not amplified the first time later did. It is possible that the quantities of trypanosome DNA in the samples were very small so when pipetting DNA from the original sample it may not always have been collected. After sequencing, only two samples gave confirmatory results for pathogenic trypanosomes, giving a total prevalence of 1.25% for pathogenic species of *Trypanosoma*. This low infection rate is consistent with a low abundance of tsetse as evidenced by the absence of any tsetse being caught in the survey traps.

It is possible that small numbers of tsetse are present in other parts of the forest reserve which the animal handlers may have visited, and it is also possible that the infected cows had been imported from other areas. The VVBDRI ranch would occasionally care for other herds if they had been confiscated and may have come from further afield. There were also large numbers of *Stomoxys* in the area, which have been shown to mechanically transmit trypanosomes (Baldacchino *et al.*, 2013) and could have spread trypanosomes if one infected animal had been brought into the herd. Several months after the cattle sampling trip the veterinarian made contact to inform about a sick cow who had been treated with trypanocides after a microscopy positive finding for trypanosomes.

If however, the trypanosome infections have been caused by local tsetse, it would show that monitoring in an area with only tsetse traps is not sufficient to declare an area tsetse-free. Barclay and Hargrove, (2005) provide an algorithm for declaring an area tsetse-free based on entomological surveys only. However, in areas where tsetse are known to have

previously infested but no longer appear to present, a sentinel herd may be useful for confirming if transmission is still ongoing. The present results provide support for such an approach.

The percentage of flies with *T. brucei* infection was higher in Saadani (2%) flies than has been reported in the Serengeti area (0.7%) (Auty *et al.*, 2012). This is surprising as the (Auty *et al.*, 2012) study examined tsetse collected from deep inside the Serengeti at the park headquarters where there would be no cattle, only wildlife. The Saadani National Park is very narrow and therefore flies will never be far from cattle or other livestock so may take less meals from wildlife, where trypanosomes are amplified. The fact that *T. vivax* had the lowest prevalence after sequencing from fly samples is a common result. It is thought that parasitaemia's are generally low and many *T. vivax* infections go undetected (Fidelis Junior *et al.*, 2019). These results confirmed that flies in the region are infected with trypanosomes and are therefore a disease risk to cattle in the surrounding areas.

4.4.6 Limitations and solutions

The logistical limitations in this chapter were similar to those in the Serengeti area. In this case, a combination of the predictive map and satellite images were used to plan where traps would be placed. The aim was to sample tsetse in areas where many tsetse were predicted to be either relatively abundant or scarce. As this was a new study area, satellite images were used to identify roads along which which traps could be placed. However, roads which looked like viable route options from Google Earth were not always accessible in practice. This was due to a lack of maintanence on roads so that vegetation had grown over the roads or trees had fallen and not been cleared, or the track had been severely eroded by rain. When this happened several traps had to be placed in new locations without exact GPS guidance so an even spread across areas predicted to have high/moderate/low catches of tsetse could not always be achieved.

For examining change over time, ideally a selection of images would have been available around a short time period to take an average value for that year but unfortunately due to the cloud cover this was not possible. Changes in Band 7 and NDVI may be complicated to examine over time outside protected areas, because it is difficult to distinguish some types of natural vegetation from growing crops which may have similar Band 7 and NDVI values.

When examining trypanosomes in cattle, only an overall herd prevalence was gained. Although the majority of cattle had ear tags, they were often on very small metal clips which were difficult to read and sometimes faded. We were however informed that very

few of the cattle in the herd had been recently imported and most were born at the farm so prevalence values should be relevant for the area.

4.5 Conclusion

The finding that NDVI and Band 7 values have not deteriorated at Mivumoni over the past 20 years suggests that non-environmental factors may have contributed to the decline in tsetse catches and the low incidence of animal trypanosomiasis. While there are many possibilities for external factors, it is likely that the use of ITC has played a large role in this decrease. Records from the ranch show an intense use of insecticides and they are commonly using ones which are capable of killing tsetse. Due to the zero grazing nature of many of the farms in the Tanga, the VVBDRI herd is much larger than others, and this number of hosts could have provided a large proportion of blood meals for tsetse in the Kibubu forest and caused a high mortality in the tsetse population here. At Pangani the environment has deteriorated but the values for NDVI and Band 7 are still comparable to those in areas which do still support tsetse so they cannot be solely to blame, and herds were present in this area which are likely to have been treated with insecticide.

The areas where the predictive models of tsetse abundance has not worked well is where it is expected that external factors have impacted on results such as at Mivumoni and Pangani. The external factors may also have led to the best predictive model for the coast being one which contains only Band 7 rather than all the environmental variables such as in the Serengeti model, as environmental effects have been diluted.

This work has shown that is it possible to take a model developed approximately 600 km away and apply it in an area which differs by climate, vegetation and elevation. Whilst the absolute predictions are much higher than would be expected, it has still shown it can be a useful tool for identifying potential areas of tsetse abundance. If it can be used in the coastal setting of Pangani district, which is markedly different to Serengeti, it shows promise for developing models to predict abundance of tsetse in other areas of northern Tanzania.

Chapter 5: Use of population genetics to examine the impact of vector control on tsetse populations

5.1 Introduction

Populations of tsetse flies across East and Southern Africa have been subjected to both natural and anthropogenic factors which have led to declines and increases in density and distribution. For instance, multiple rinderpest epizootics since the 1890s caused a huge decrease in densities of tsetse host species and subsequently a decrease in tsetse population numbers (Leak, 1998). In countries such as Kenya there has also been large scale control programmes following the introduction and use of insecticides. Changes in land use from low intensity pastoralism to high intensity crop farming and/or deforestation for fuel have also reduced the numbers of possible habitats and fragmented remaining ones, as well as reducing the numbers of wild hosts presents. These events have in some areas caused genetic bottlenecks which can be studied to determine if and how the populations have declined and, possibly, rebounded. It is also possible to determine the geographical direction of gene flow. Such knowledge may assist the design and implementation of control operations, since it will indicate the likely sources and direction of tsetse populations re-invading an area subject to tsetse control. Population genetics methods will be discussed here in terms of applicability to examine possible ongoing control at the Serengeti study sites (Chapter 2 and 3), and at the Tanga sites (Chapter 4), over short (~10 – 30 km) and longer distances (~600 km). Literature describes previous studies which have been conducted across several tsetse species and across habitats of varying distance from each other, so could be a useful tool to quantify the possible impact of control around the Serengeti National Park and in Tanga region.

5.1.1 Methods and measures of population genetics

Many of the first population genetics studies were conducted with allozyme loci. An allozyme is a different form of an enzyme which is not functionally different, but is structurally different from another allozyme coded for by a different allele at the same loci (Van Etten, 2011). Allozymes however have been shown to overestimate gene diversity as heterozygotes are more likely to survive through natural selection (Krafsur, 2002a). Later studies have been conducted with mitochondrial markers; mitochondrial DNA is located in the mitochondria of cells, organelles responsible for producing the cell's energy. For population genetics studies single nucleotide mutations between individuals can be examined. Mutations occur relatively frequently in mitochondrial DNA as it lacks the proof-

reading capacity of nuclear DNA when replicating (Krafsur *et al.*, 2001). More recently, microsatellite markers have been used in both small and large-scale studies. Microsatellites are single sequence repeats, normally of 2-5 base pair repeats such as ATATAT or CAGCAGCAG which are present across the nuclear genome. They are useful for population genetics studies as they are highly polymorphic (Solano *et al.*, 1999), and are rarely subject to selection pressure due to the phenotypic advantages they confer (Vieira *et al.*, 2016).

One of the most common ways to examine genetic variation is through gene diversity, otherwise known as expected heterozygosity (He). This is the predicted probability that for the locus in question, the individual will be heterozygotic. Low gene diversity would be expected in very small populations and a high expected heterozygosity means there is a high level of genetic variability. A gene diversity of 0.6 would indicate a 60% chance of the individual being heterozygotic. Related to this, allelic richness defines the actual number of alleles within a population at a specific locus and is a measure of how well a population may fare when subjected to long-term selection pressure, with a larger value being better.

One of the first fixation indices, or F statistics, which can be examined is F_{is}. This calculates the average excess of deficit of homozygotes within a subpopulation. An F_{is} score is given between -1 and +1, with a score closer to +1 indicating a more inbred group, however positive F_{is} scores can also indicate technical problems with the selected markers. The P values for an F_{is} score determine if a locus is significantly deviating from Hardy-Weinberg equilibrium. The Hardy-Weinberg principle relates to a population where there is no selection, no mutations, no migration, and no non-random mating which results in the genetic diversity of a population being maintained with an equilibrium of the genotype. Another commonly examined factor is genetic differentiation. Genetic differentiation describes how related populations are to one another and is given by F_{st}. A large F_{st} value would describe two populations which are highly differentiated, with very little mating between them and therefore very low gene flow. Geographical distance can play a large part in genetic differentiation, although it depends on whether there is continuous suitable habitat enabling greater movement, and the species of fly involved.

When designing genetics studies, each marker should be independently inherited and not linked to any others as this could introduce bias. Linkage disequilibrium describes the association of alleles at loci which are not independent in their inheritance. A P value can be given for each pair of loci to determine if they are significantly linked across the populations.

Studies to identify genetic bottlenecks have been carried out with several tsetse populations from East and Southern Africa. A population bottleneck is an event which describes a sudden and significant reduction in the size of a population. This results in a genetic bottleneck, where the genetic variation within a population is reduced as the number of mating individuals are reduced.

5.1.1.1 Allozyme markers

Krafsur *et al.*, (1997) originally used eight allozyme loci to examine gene diversity and gene flow of east and southern African *G. pallidipes,* which included Kenya, Mozambique and Zimbabwe. Genetic diversity was higher in the southern African countries than in Kenya, although not significantly. It was expected that populations across different countries could be genetically close as the species is continuously distributed, however the results showed a high level of genetic separation, with genetic drift being a strong factor.

5.1.1.2 Mitochondrial markers

Krafsur and Wohlford, (1999) examined population structure of *G. pallidipes* from Zimbabwe, Mozambique, Kenya and Ethiopia using four mitochondrial loci. Previous studies with allozyme loci had indicated a high level of population structuring which was contradictory to ecological studies. Their results found high levels of differentiation in Ethiopia and Kenya compared to Zimbabwe and Zambia where variation and differentiation was much lower. They concluded that genetic drift was responsible for the high levels of differentiation as populations ended up scattered following the rinderpest outbreak. All samples which were within 30 km of each other clustered together in a genetic dendrogram. Between countries there was one well defined cluster which included Zimbabwe and Zambia which is unsurprising as they are geographically closer to each other, and there is a vegetation belt of relatively suitable habitat between the two areas. These results show that within 30 km these mitochondrial loci may not be suitable for analysis, but this clustering may be less pronounced over the same distances in countries north of Zambia where variation is generally greater.

Several large scale studies have been conducted to examine genetic diversity of tsetse across Africa. Krafsur, Marquez and Ouma, (2016) used a 300 bp fragment from a ribosomal (*r16S2)* gene and a 440 bp fragment from the cytochrome oxidase I (*COI)* mitochondrial gene to examine *G. pallidipes* population genetics across east and southern Africa. Mitochondrial variation is more appropriate for detecting demographic forces than microsatellite variation and therefore more suited for geographically larger-scale studies. Their study found Ethiopian fly populations to be the most genetically diverse and

phylogenetically the oldest, in comparison with southern African (Zambia and Zimbabwe) flies which were the least genetically diverse and phylogenetically youngest. This they argued, was linked to rinderpest outbreaks in southern Africa which caused population suppression of tsetse, whereas Ethiopia was not as heavily affected by the epizootic. Diversity in Tanzania was found to be greater than in Zambia and Zimbabwe which correlates with similar microsatellite studies (Ouma, Marquez and Krafsur, 2005) and should provide stronger signals within the country.

5.1.1.3 Microsatellite markers

Three microsatellite loci were examined along with two mitochondrial loci and eight allozyme gene loci (Krafsur, 2002b) from flies collected in Kenya, Zimbabwe and Mozambique. Microsatellite and mitochondrial diversity was significantly higher in Kenya than it was in Zimbabwe and Mozambique, credited in part to the genetic bottleneck from the rinderpest outbreak. As Tanzania is north of Mozambique and next to Kenya, there is the possibility that it too was less genetically affected by the outbreak. Allozyme diversities showed the opposite effect, with greater diversity found in samples from Zimbabwe and Mozambique than Kenya. Allozyme analysis did not detect a bottleneck in Zimbabwe or Mozambique, which shows that it would not be useful for the current planned study in Tanzania which aims to detect small-scale bottlenecks.

Krafsur and Endsley, (2002) state that before an area-wide control programme can begin, it is important to understand the gene flow of the flies in that area. They used six microsatellite loci for seven *G. m. submorsitans* populations in The Gambia and Ethiopia and five microsatellite loci in six *G. m. morsitans* populations from Zambia, Mozambique and Zimbabwe. With the exceptions between Mana, Angwa and Makuti which are approximately 100 km apart, there was very little gene flow between all areas. The gene diversity was high and there was no evidence of previous major bottleneck effects, which suggests many flies persisted undetected through the rinderpest outbreak, as it is unlikely the mutation rate would have been so high to have the genetic diversity completely recover in approximately eight generations.

Ecological studies have shown tsetse moving larger distances and populations being more dispersed than that suggested by genetic studies (Gooding and Krafsur, 2005). Ouma, Marquez and Krafsur, (2005) aimed to increase the body of evidence to support genetic results by increasing the number of microsatellite loci which had been used in large geographic studies of *G. pallidipes* from three to eight. The authors also used five mitochondrial markers to add weight to the results. A total of 720 flies from 21 populations

in six countries were analysed. A high level of genetic differentiation was observed, even between populations only tens of km apart, suggesting low rates of gene flow and movement.

Nguruman and Lambwe valleys in Kenya provide an example of areas where tsetse have previously been subjected to high levels of apparently successful control but have since recovered. Ouma, Marquez and Krafsur, (2006) used eight microsatellite loci to examine whether the current populations were mainly endemic or if immigration had repopulated the area. They used traps which ranged from 200 m to 14 km apart, with 15 sites at Lambwe grouped into three blocks and 27 sites at Nguruman grouped into six blocks, and also took samples from Kodera (50 km west of Lambwe) and the Serengeti (170 km from Nguruman) as the closest known tsetse infested areas. Within each site there was no genetic differentiation, however each block was significantly differentiated from each other. The results suggested that Lambwe and Nguruman are geographically isolated from Kodera and the Serengeti, and that populations currently present are highly endemic. This paper would indicate that even over several km, a geographic structuring can be seen.

These findings are partially corroborated through ecological surveys. Between 1993 and 1994 in Nguruman, Odulaja *et al.*, (2001) collected *G. pallidipes* monthly at 20 trap sites, with each adjacent site approximately 800 to 2300 m apart. During the wet season, density of catches was even across the area which suggested a random distribution – however, during the dry season the catches were aggregated and positive correlations between traps was only significant when the distance was less than 3.8 km.

Okeyo *et al.*, (2017) also examined *G. pallidipes* in two regions of Kenya for temporal and spatial genetic differentiation. The two areas; Ruma National Park, and Nguruman Escarpment, were both subject to a high level of tsetse control between 1960 and 1980, while Nguruman was also subjected to additional habitat change caused by humans. Samples were collected between 2003 and 2015 and within each area three trapping sites were used which were approximately 20 km from each other. There was no significant geographical structuring found at either site, with no correlation between genetic and geographical distances from samples collected during the same year. This is at odds with the results of Ouma, Marquez and Krafsur, (2006) and while this is acknowledged in the paper, suggestions as to why this has occurred are not offered. The authors compared this with comparable distances for *G. f. fuscipes* which did show significant differences in genetic variation at comparable scales and suggested that the geographical environment

has allowed for greater gene flow in their study areas or that *G. pallidipes* is covering greater distances in comparison to what *G. f. fuscipes* would. However, a strong temporal variation was seen at Nguruman but less so at Ruma. Genetic bottlenecks were not detected at either site which indicates the high level of tsetse control carried out in earlier decades has had little lasting impact on genetic diversity, however it is thought that the modification of habitats in Nguruman by human use in more recent years could have caused the temporal genetic changes which are seen more strongly. This again was at odds with previous studies conducted in the same area which found no temporal change, however, for this the authors points out that Ouma, Marquez and Krafsur, (2006) covered only ~ 2 - 8 generations, whereas this study covered ~ 96 generations. In this study 13 microsatellite loci were used which has proved sufficient for examining temporal change over 12 years but not spatial variation. This suggests a greater number of microsatellite loci may be needed to detect lower levels of variation over these distances.

Ciosi, Masiga and Turner, (2014) examined genetic bottlenecks within the International Atomic Energy Agency (IAEA) colony of *G. pallidipes* flies and attempted to trace the origins of the population. They used 9 microsatellite loci and performed quantitative analysis using Approximate Bayesian Computation. They determined that the establishment of the colony was associated with a loss of 35.7% of the alleles compared to the original source population. They also showed a possible effective population size reduction. Their examination of the IAEA colony was due to concerns regarding the use of colony flies for SIT as differentiation may lead to issues with mating and colony flies will have had a large selection pressure to adapt for laboratory conditions making them less suitable for the field.

A previous study by Manangwa *et al.*, (2019) has recently used microsatellite data to examine relationships between *G. pallidipes* populations at seven sites in the Serengeti area. After confirming the species of the flies through morphological identification, 10 microsatellite loci were selected. Four of the loci failed to amplify for almost a third of the 190 samples and were removed from the analysis. Tests of linkage disequilibrium found strong effects, with each of the six loci found to be in a minimum of two significantly linked pair of loci. There was a significant homozygote excess and null alleles were found at each locus. The authors considered that a Wahlund effect was occurring, where genetically distinct populations are pooled together resulting in an apparent reduction in heterozygosity. After analysing the samples in BAPS, they discovered five clusters of samples. The samples were then split into these new clusters and reanalysed. From this

they calculated new Cavalli-Sforza chord genetic distances to create a Neighbour Joining Tree. This showed two separate lineages; one with 26 clusters and the other with seven clusters. When the samples were re-analysed and split into these new clusters, the locus pairs were no longer in a significant linkage disequilibrium, and the homozygote excess turned into a slight heterozygote excess in one clade, with no significant excess in the other. As this paper examines *G. pallidipes* in Tanzania and specifically in one of the same study areas as the present study, it seems likely that these initially confusing results may also be found in this study. In this study the aim is to analyse tsetse from approximately 30 flies collected from each site as the Manangwa *et al.*, (2019) study has done, so if populations which are not panmictic are present they should be detected.

5.1.2 Chapter objectives

It appears that population genetics studies on *G. pallidipes* may have mixed success at determining structuring or control. In some instances, a genetic structuring was detected in populations separated by just tens of km, and in others, differences could not be found over hundreds of km. There could however be many factors which had caused differences between studies such as the amount of vegetation which was linking sites of similar distances, the season which they were collected in, and the type and number of marker they had used. It is likely that determining population structures is easier with riverine species of flies, but this is because they are ecologically less mobile. Microsatellite markers should be useful for examining populations over short distances due to their highly polymorphic nature, and if enough markers can be used this should increase the ability to detect genetic structuring.

Work conducted in Tanga region (Chapter 4) showed that the numbers of tsetse have been reduced to undetectable levels in farming areas such as Mivumoni and Pangani where tsetse were historically caught in abundance and AAT was a major cause of livestock mortality and morbidity. Tsetse were however found in Mgambo and Saadani, protected conservation areas which are about 20 km apart. Prior to increases in human populations and associated farming-related declines in natural vegetation, it is highly likely that these two areas were connected by a continuous belt of habitat, that covered much of East and Southern Africa. Across the coastal region of Tanzania, habitat suitable for tsetse is likely to have decreased as human populations have increased, although there still appears to be a substantial level of suitable tsetse habitat between the two areas from satellite images (Chapter 4). However, if as predicted from results around Pangani that insecticide use has been ongoing at high levels, it is possible that the fly populations may have become genetically isolated from each other in the last 20 years due to what is effectively a barrier of control.

In contrast to predicted decline and genetic isolation of tsetse across much of Tanga region, the fly populations inside Saadani and Mgambo may have increased. The Saadani population was subjected to high levels of control between the late 1980s and 2000 when the site was a commercial ranch. The Mgambo Forest Reserve was not a designated rangeland where cattle could graze and the dense vegetation within it would deter livestock keepers and their cattle. The reserve, as the name suggests is also a designated conservation area with some legal protection from settlement and exploitation. It therefore seems likely that the tsetse population inside Mgambo would have been subjected to lower

levels of vector control. Analyses of remote sensing data (Chapter 4), showed that within Saadani, vegetation has become more suitable for flies since 1997 due to the removal of cattle.

There is a different historical perspective to the areas surrounding the Serengeti National Park. While the rinderpest epizootic would have affected this population, anthropogenic changes to vegetation and tsetse control operations were limited within the park. Outside the Serengeti National Park, agriculture and large-scale treatment of cattle with insecticide have had a marked effect on the tsetse population. The genetics of tsetse populations at the centre of the Serengeti National Park or outside may differ, reflecting the impact of these natural and anthropogenic differences. Population genetics methods can be used to examine the relatedness of populations and determine if the populations have been subject to any severe decline in numbers.

The aims of this chapter are to determine:

- Are Mgambo and Saadani genetically isolated? They are not geographically far apart from each other, but if the use of insecticide treated cattle between the sites is as high as predicted they may have become genetically separated.
- Have tsetse populations recovered in Saadani from the high levels of control when it was an operational ranch?
- Can the impact of vector control be assessed through analysis of potential genetic bottlenecks?
- Are populations on the edge of the Serengeti and neighbouring game reserves related to those in the middle of the Serengeti, or can they be distinguished from one another?

5.2 Methods

5.2.1 Field methods

Tsetse flies were collected from two Regions of Tanzania; Mara and Tanga (Figure 62). The Tanga and Mara trap sites are almost 600 km apart. All tsetse were collected using Nzi traps baited with acetone (500 mg/h), 1-octen-3-ol (0.4 mg/h), 4-methylphenol (0.8 mg/h) and 3-n-propyphenol (0.1 mg/h) dispensed using the methods of Torr *et al.* (1997). Flies would be collected from the traps every 24 h and stored in individual tubes containing 99% ethanol.

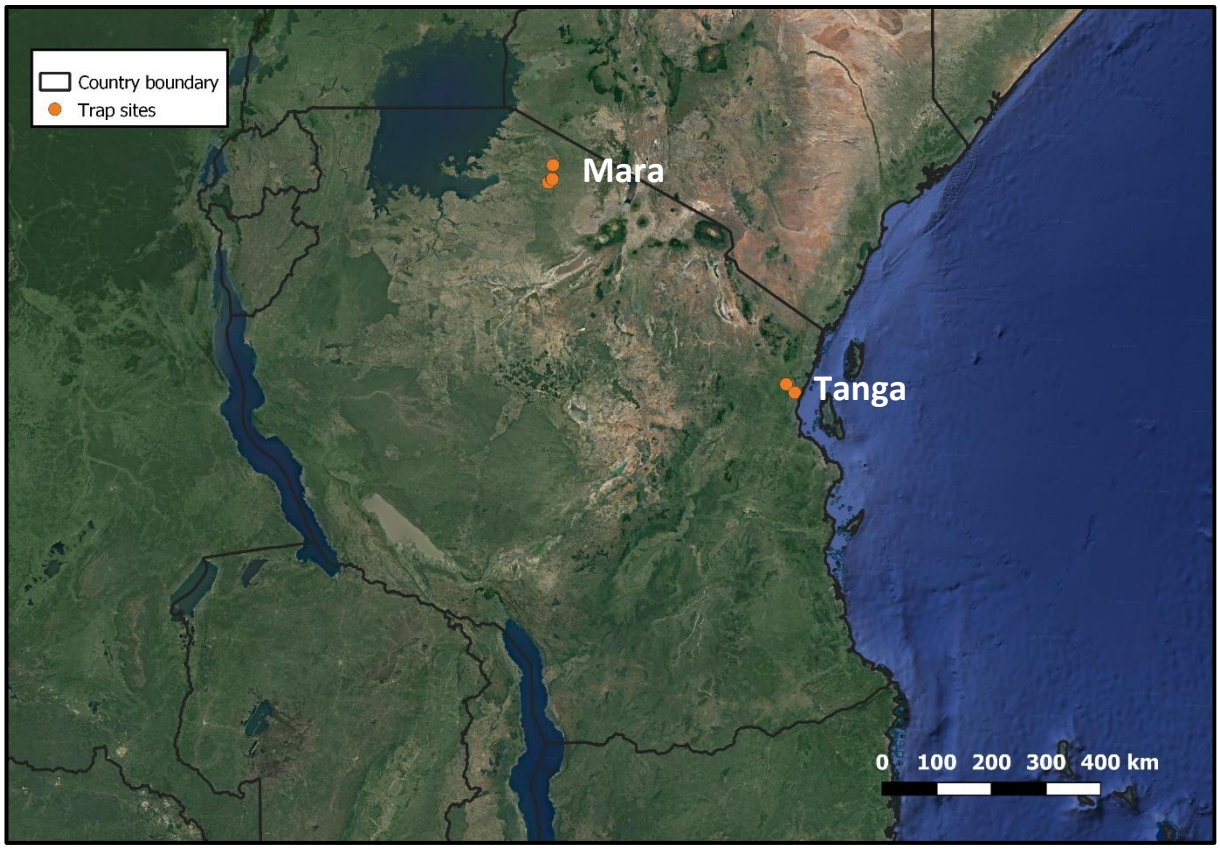

Figure 62. Location of trap sites across Tanzania.

In Mara Region, flies were selected from three separate sites (Figure 63). The flies from Ikorongo were collected over two days in July 2016 from two traps placed 300 m apart. The flies from Ikoma were also collected over two days in July 2016 but to gain sufficient samples they were taken from five traps, the furthest two being 2.2 km apart. The flies from the Serengeti were collected over two days in July 2016 from two traps placed 100 m apart. The distance between the Serengeti traps and the Ikoma traps was approximately 10 km, and from the Serengeti to Ikorongo traps was approximately 33 km. Between the Ikorongo traps and the Ikoma traps was approximately 25 km. Sampling strategy was not designed specifically for this work. All of the flies collected from the Serengeti were caught

as part of the BBSRC ZELS project to examine abundance, trypanosome prevalence and bloodmeal composition.

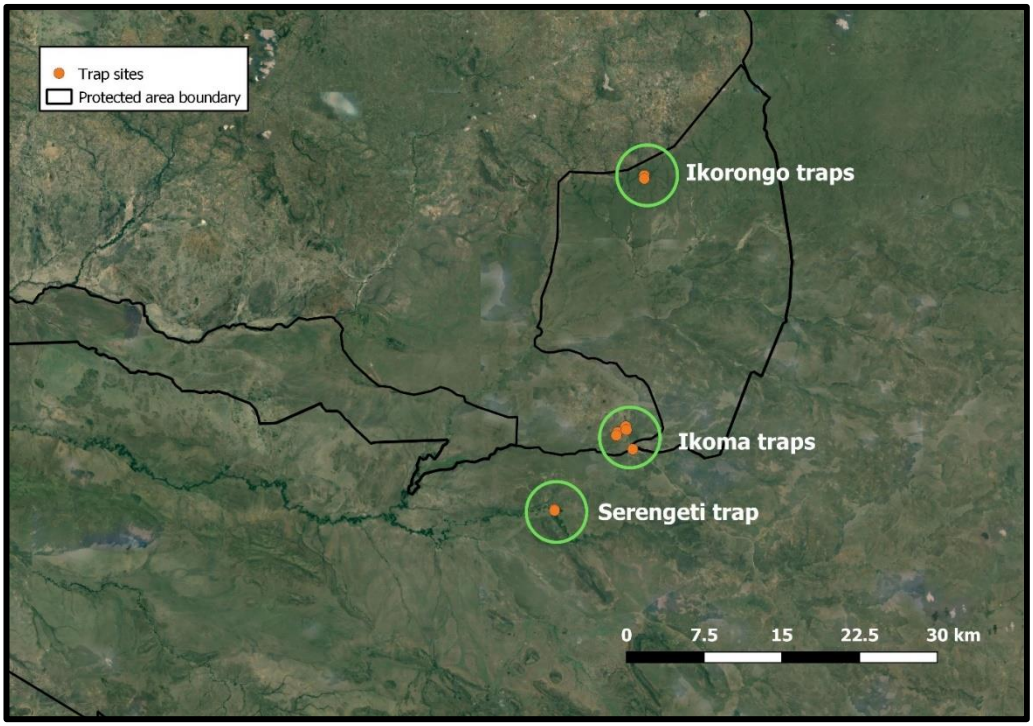

Figure 63. Mara trap sites.

In Tanga Region, tsetse were selected from two separate sites (Figure 64). The tsetse from Mgambo were collected over three days in September 2017 from two traps placed 600 m apart. The flies from Saadani were collected over three days in February 2018 from one trap. The Mgambo and Saadani traps are approximately 20 km apart. Again, the sampling strategy was not designed specifically for this work. The flies caught in Tanga region were caught to examine abundance and trypanosome prevalence (Chapter 4).

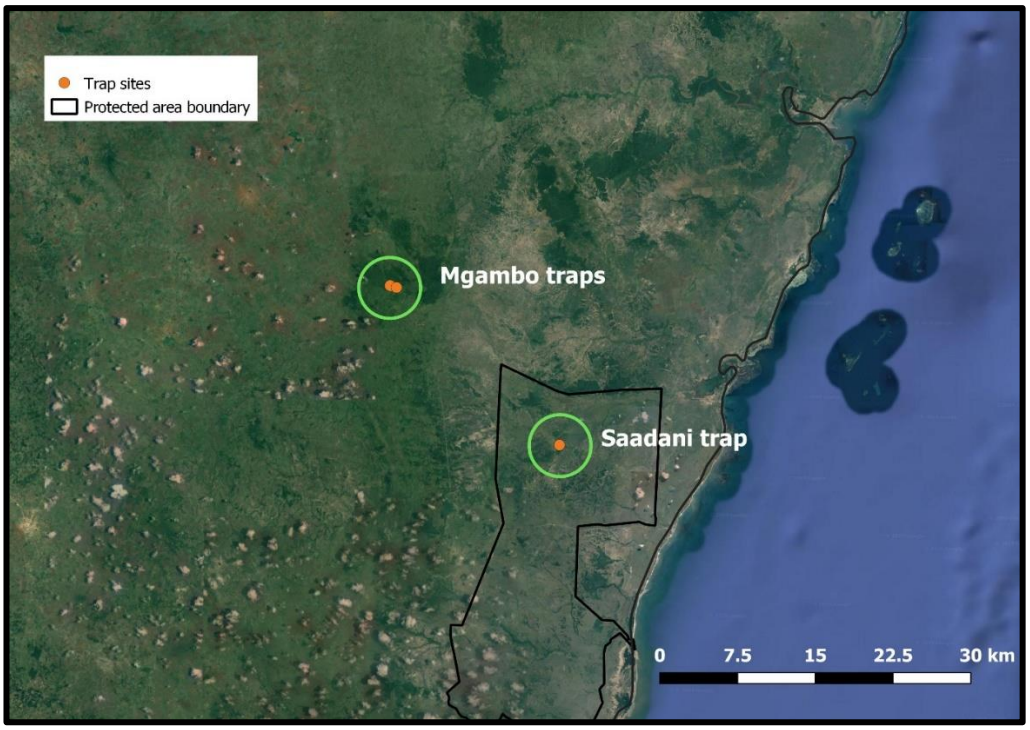

Figure 64. Trap sites in Tanga Region.

A minimum of 30 flies were collected from each site. Only males were used as the Mara flies had already been extracted for another study and had been crushed entirely so DNA from a larva could have been in any female samples.

5.2.2 DNA extraction

Flies were stored in ethanol which was first evaporated using a heat block set to 56 °C for several hours. They were then placed in fresh individual tubes each with a steel ball. A mix was then made for digestion (Table 22) from the GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific), of which 200 µl was added to each fly tube.

Table 22. Quantities for digestion solution.

	X1
Digestion Solution	180 µl
Proteinase K	20 µl

Tubes were then securely sealed and placed into the tissuelyser. The tissuelyser was set to level three for two min. Tubes were then removed and spun down in a centrifuge at 1500 rpm for 30 s, before being placed in an incubator at 56 °C overnight with the lids sealed on.

The next morning samples were removed from the incubator and spun again in the centrifuge at 1500 rpm for 30 s. Samples were then each transferred to a new Eppendorf and 20 µl of RNase A solution added to each tube, before being vortexed and left at room temperature for ten min. Then 200 µl of lysis solution was added to each sample and vortexed for 15 s. After this 400 µl of ethanol was added, vortexed again and spun briefly in the centrifuge.

The prepared lysate was then added to a GeneJet Genomic DNA Purification Column into a collection tube. The collection tubes with the column in were then centrifuged for 1 min at 6000 x g RCF. The GeneJet column was then placed in a new collection tube and the solution in the old collection tube was discarded. 500 µl of wash buffer 1 was then added to each sample and centrifuged for 1 min at 8000 x g (RCF). The flow-through from this was discarded and the purification column placed back into the collection tube. 500 µl of wash buffer 2 was then added to each column and centrifuged for 3 min at 12000 x g (RCF). The GeneJet column was then transferred to a sterile 1.5 mL microcentrifuge tube. 200 µl of elution buffer was then added to the centre of the GeneJet column membrane to elute the genomic DNA. Samples were then incubated for 2 min and then further centrifuged for 1 min at 8000 x g (RCF). 200 µl of this extracted DNA was transferred to a new PCR plate for storage.

5.2.3 Primer and multiplex design

To design primers specific for *G. pallidipes* a file of contigs was downloaded from VectorBase. The sequences belonged to the gene set Gpall 1.7 which is from a colony maintained by the IAEA laboratories in Austria.

The files were unzipped in FA format and imported into SciRoKo (Kofler, Schlötterer and Lelley, 2007). In the search mode settings of SciRoKo, the selected options were perfect repeats with a minimum number of repeats set at eight. In the simple sequence repeats (SSR) stats setting box the upper and lower boundary were both set at 4, with the minimum required score and minimum count left at the default. The search was done and in display statistics it was set to export the SSRs as tab deliminated.

This file was then opened in excel and filtered for sequences with 10-14 repeats. An expression was used to highlight any sequences which were in the same contig and duplicates removed. Flanks for the remaining contigs in the list were searched for using SciRoKos's little helper with 200 bp set as the flanking length on each side. The list was then filtered further to remove any AT or CG rich repeats and any sequences which had over four

single base repeats or four double base repeats. The sequences were also checked to ensure there were no repeat regions in the flanking areas.

This gave a list of 60 possible sequences. These sequences were run in NCBI primer blast (Ye *et al.*, 2012) with the data set selected as *G. pallidipes* and a desired length between 80 – 300 bp. Primers which gave an annealing temperature close to 60 °C and had values less than 3 for "self 3' complementary" were selected, with an A set and a B set in case the first set did not work. The selected primers were run in NCBI blast (Altschul *et al.*, 1990) to see if they would match with any non-target sequences.

After narrowing the list down by the above criteria, this gave a total of 40 primer sets. These were then loaded into Multiplex Manager (Holleley and Geerts, 2009). Based on the annealing temperatures and product size, they are sorted into a selection of multiplexes and assigned one of four coloured labels; 6FAM (blue), VIC (green), NED (yellow), PET (red). The one with the fewest overlapping product sizes was selected which had five multiplexes with 6 primer sets and two with five primer sets. These primers were then ordered from Macrogen.

5.2.4 Laboratory methods

The labelled forward primers were diluted with buffer (1mM Tris-HCL pH 8.0/1 mM EDTA). All of the forward primers were of 10 nmol concentration and were diluted with 100 µl of buffer. The reverse primers were diluted with nuclease free water to a 100 µM concentration. Once suspended the primers were stored in the -20 °C freezer, with special care taken to ensure the labelled forward primers were covered and only exposed to minimal light whilst being diluted.

For each plex a master primer solution was made up. The concentration for each primer was 2 µM so 10 µl of each forward and each reverse primer for each plex was pipetted into a tube. For the multiplexes containing 6 primer pairs, this gave a total primer volume of 120 µl, which was added to 380 µl of water to give a total solution of 500 µl. For the multiplexes containing 5 primer pairs this gave a total primer volume of 100 µl, which was added to 400 µl of water to give a total solution volume of 500 µl. A mastermix was then made up based on 2 µl of sample in each well (Table 23).

Table 23. PCR mastermix components.

Reaction component for multiplex PCR	X1
2X Type-it PCR MM	6.25 ul
10X primer mix	1.25 ul
Water	3 ul
Template DNA	2 ul

The PCR protocol used is as follows:

	Time	Temperature	
Initial activation	5 min	95	
Denaturation	30 s	95	
Annealing	90 s	59	30X
Extension	30 s	72	
Final extension	30 min	60	

To confirm the presence of a PCR product the samples were run on a gel. Eighty mL of TAE and 0.8 g of agarose were combined in a flask and microwaved for 1 min to dissolve the agarose. Ten µl of Peqgreen was then added to the flask. The contents were poured into a tray and left to set with well combs in. Once the gels were set they were placed into the tanks and the tanks filled with TAE. Five µl of sample was combined with 1 µl of 5X loading dye and placed into each well. Five µl of DNA ladder was placed into the first and last well of each row. Gels were run for 65 min at 105 volts. Once the gel has finished running it was drained and placed into the gel imager. The programme GeneSnap (SynGene) was used to visualise the DNA in the gel to confirm if products had been produced and that the negative control had not produced a band. To prepare samples for shipping, 2 µl of sample was added to 18 µl of nuclease free water and these were then sent to Macrogen Europe.

Initially only 12 samples were sent for microsatellite analysis to check the balance of the primers within each plex. The results were satisfactory, and the remaining samples were processed using the same primer concentrations.

5.2.5 Sequencing

A subsample of flies (31) were also sent for sequencing using mitochondrial DNA primers. The primers used were for a 316 bp section of 16S region (Dyer *et al.*, 2008). The mastermix was created (Table 24), and 24 µl of mastermix was added to each well with 1 µl of sample.

Table 24. 16S primer mastermix.

Component	X1
Water	9.5 µl
Forward primer	1 µl
Reverse primer	1 µl
Taq	12.5 µl

The PCR protocol for this was as follows:

	Time	Temperature (°C)	
Initial activation	5 min	95	
Denaturation	30 s	95	
Annealing	60 s	57	34X
Extension	60 s	72	
Final extension	10 min	72	

The PCR products were run on a gel as above (section 5.2.4), and once presence of a product had been confirmed through gel imaging the samples were prepared for sequencing. To do this, 0.5 µl of Exo I (Exonuclease I), and 1 µl of rSAP (Shrimp Alkaline Phosphatase) was added to 5 µl of PCR product. The mix was incubated at 37 °C for 15 min then at 80 °C for 15 min. Samples were then sent for sequencing with Source Bioscience (Rochdale, UK).

The flies were also sequenced with AltCOI mitochondrial DNA primers (Cunningham, unpublished) which targeted the COI region and gave a product size of 1040 bp (F: 5'-ATG GWG GAG CTT CAG TWG ATT TAG-3', R: 5'-TAA TCA TTC AAT WGA AGA ATT TAA TTG AAT AGG-3'). PCR protocol was the same as for the 16S primers but the annealing temperature was 55 °C.

5.2.6 Analysis methods

GeneMarker V3.0.0 was used to call the allele sizes with a size standard of 500LIZ. The template for microsatellite dinucleotide was initially used to run each plex. Then the panel editor was opened and modified with the ploidy set as tetraploid. Each locus was assigned its colour and the expected product size drawn on.

The samples were then re-run with the modified template. The allele call range was set at 80-400 and a peak detection minimum intensity of 1000 and a maximum intensity of 60,000. Stutter peak and plus A filters were applied. Each label within each plex was analysed individually and all uncertain peaks were called manually or discarded. Minor scoring errors can occur so peaks which had been called out by 1 bp were adjusted.

The final calls were then exported to Microsoft Excel. Data from each plex was then combined and the MS Tools add in was used to firstly error-check the data. The input data format was 'Diploid – Two-column format' and a minimum and maximum expected allele size were given as 80 and 400. It was also used to check for jumps in allele size over 40. After error checking, the data was converted from diploid two-column to diploid one column and converted into input forms for FSTAT and Genepop (3-digit).

FSTAT was used to analyse the genetic structure of the five populations (Goudet, 1995). Through FSTAT the tests completed gave gene diversity per locus and population, allelic richness per locus and population and F_{is}. Genepop on the web (Raymond and Rousset, 1995) was used to test for linkage disequilibrium and F_{st}. The parameters for the LD test were genotypic disequilibrium using the log likelihood ratio statistic, with default Markov chain parameters.

To make a formal test of diversity the comparison among groups of samples tab in FSTAT was used. A two-sided test for two groups was applied. All boxes were ticked to include statistics to compare among groups – Allelic richness, obs. Heterozygosity HO, Gene diversity HS, F_{is}, F_{st}, relatedness and corrected relatedness. Two groups were selected with the three Serengeti sites in one group and the two coastal sites in the other group.

The programme Bottleneck (Cornuet and Luikart, 1996) was used to compare heterozygosity from the observed allele frequencies with those expected at mutation-drift equilibrium under particular mutation models. This can suggest if a genetic bottleneck or an expansion may have occurred through a Wilcoxon two-tailed test. The parameters selected were a two-phase model which was run with a frequency of larger (non-single step) mutations of 0.3.

To estimate relationships between flies within and between populations, ML-Relate (Kalinowski, Wagner and Taper, 2006) was used. The data was imported in genepop format and the relationships output for which flies appear genetically unrelated, like half-siblings, full-siblings or have a parent-offspring sharing of alleles shown in a matrix. To calculate effective population sizes (Ne) a program called NeEstimator (Do *et al.*, 2014) was used.

After this, BAPS (Corander *et al.*, 2008) was used to analyse for clustering of individuals. The data were entered in genepop format and the simulation was run with the upper bound to the number of populations run with 5, 10, 15 and 20 populations which was run three times. This creates an output which states which number of clusters is significantly most likely to be real. BAPS admixture based on mixture clustering was then used with parameters of 5 for the minimum size of a population that will be taken into account when admixture is estimated. For the input number of iterations this was 100, and the input number of reference individuals from each population was set at 200 and the input number of iterations for reference individuals was set at 10.

To quantify the bias caused by null alleles and whether they could still be used for structural analysis, the software FreeNA was used (Chapuis and Estoup, 2007). This software estimates what the value of the other allele would have been if present and estimates the expected frequency of it. This can then be compared against the null allele to determine the impact on genetic diversity. For this, replicates were set at 1000.

The output from BAPS gave an indication for the number of possible clusters to test in STRUCTURE (Pritchard, Stephens and Donnelly, 2000). The genepop files were first converted in STRUCTURE format with PGDSpider (Lischer and Excoffier, 2012) then loaded in a new STRUCTURE project. The parameters used were a burn-in period of 50,000 with an MCMC reps of 100,000. The ancestry model used was an admixture model, the allele frequency model was allele frequencies correlated, and 'compute probability of the data (for estimating K)' was selected. All five populations were initially run together for possible K 2 to 9, for 9 iterations. The Tanga populations were also run without the Mara ones to determine any further structuring. The output results folder was then zipped and run through STRUCTURE harvester (Earl and vonHoldt, 2012) which suggests the number of K (the most likely number of genetic groups). These results can also be visualised in CLUMPAK (Kopelman *et al.*, 2015).

The populations were then re-grouped based on the BAPS and STRUCTURE results. Tests were repeated for allelic richness, F_{is}, F_{st}, LD, Ne and bottlenecks for the new populations.

Genetic distance between these new populations was calculated using GENDIST within PHYLIP (Felsenstein, 1989) using the Cavalli-Sforza chord measure. The data was converted from genepop format into PHYLIP format using Convert 1.31, then the PHYLIP output was converted using PGDSpider into MEGA format. The genetic distance matrix was used to create a Neighbour Joining Tree within the Phylogeny section of MEGA (Tamura *et al.*, 2013). MEGA was also used to create a Maximum-Likelihood Tree with 1000 bootstraps for the AltCOI mitochondrial sequences.

The 16S sequencing results were then examined using NCBI blast to determine cryptic species structuring. Finally, to confirm if the loci used in the study were capable of successfully clustering cryptic species the software WHICHLOCI (Banks, Eichert and Olsen, 2003) was used.

5.3 Results

5.3.1 Sample sizes post data cleaning

A total of 153 samples were sent for sequencing using 40 loci. After removing samples and loci where 25% of data was missing, this left 145 samples from 32 loci (Table 25).

Table 25. Sample sizes for each population post data cleaning.

Population	N
Ikoma	23
Ikorongo	30
Serengeti	29
Mgambo	33
Saadani	30

5.3.2 Gene diversity

The area with the largest gene diversity across all areas was Mgambo followed by Saadani. From the Mara sites the population with the largest gene diversity was the Serengeti with Ikoma as the lowest. Within each region, the gene diversity was similar (Table 26).

Table 26. Gene diversity across the five populations with 32 loci.

Gene diversity per locus and population					
Loci	**Ikoma**	**Ikorongo**	**Serengeti**	**Mgambo**	**Saadani**
1052A	0.701	0.707	0.778	0.601	0.357
1383A	0.735	0.783	0.542	0.853	0.864
2021A	0.173	0.131	0.219	0.476	0.352
0027A	0.133	0.186	0.287	0.71	0.586
0702A	0.409	0.509	0.355	0.595	0.591
2107A	0.127	0.413	0.336	0.389	0.516
3840A	0.638	0.632	0.559	0.773	0.782
1333A	0.381	0.454	0.39	0.452	0.512
1445A	0.496	0.514	0.543	0.684	0.726
0900A	0.675	0.66	0.646	0.809	0.728
5042A	0.855	0.807	0.821	0.734	0.677
0483A	0.444	0.394	0.37	0.737	0.767

0105A	0.128	0.298	0.397	0.545	0.498
2048A	0.501	0.402	0.499	0.646	0.63
0065A	0.713	0.608	0.744	0.703	0.737
4844A	0.126	0.033	0.068	0.545	0.59
0138A	0.699	0.715	0.714	0.901	0.928
2858A	0.342	0.067	0.256	0.559	0.599
0585A	0.63	0.71	0.667	0.84	0.793
0715A	0.091	0.251	0.259	0.831	0.803
0649A	0.499	0.49	0.372	0.504	0.428
4108A	0.575	0.481	0.623	0.499	0.461
0243A	0.645	0.62	0.641	0.761	0.676
0064A	0.664	0.658	0.759	0.764	0.846
0026A	0	0	0	0	0
2425A	0.717	0.673	0.692	0.6	0.649
1926A	0.551	0.633	0.591	0.858	0.773
1388A	0.312	0.191	0.361	0.424	0.369
2434A	0.128	0.297	0.227	0.496	0.675
5004A	0.873	0.854	0.76	0.696	0.73
0264A	0.298	0.186	0.262	0.579	0.729
0752A	0.634	0.676	0.602	0.55	0.672
Average	0.465406	0.469781	0.479375	0.628563	0.626375

5.3.3 Allelic Richness

Allelic richness to examine diversity within populations puts Saadani highest with the largest value for average allelic richness, followed by Mgambo, Ikoma, Serengeti and lastly Ikorongo. Again, values from the two Tanga sites were similar to each other, and values from the Mara sites were comparable with each other (Table 27).

Table 27. Allelic richness across the five populations with 32 loci.

Allelic Richness per locus and population based on a minimum sample size of 21 diploid individuals					
Loci	**Ikoma**	**Ikorongo**	**Serengeti**	**Mgambo**	**Saadani**
1052A	6.82	6.29	6.919	8.512	5.301
1383A	4.998	6.527	4.652	8.723	9.434

2021A	2	2.827	2	3.655	5
0027A	3.864	2	3.905	4.871	5.632
0702A	3	3.7	3.5	3.635	3
2107A	3.739	3.89	3.963	4.271	4.67
3840A	3	3.914	2.996	6.808	5.676
1333A	2.913	3.952	3.448	2.956	3.4
1445A	2	2	3.855	3.999	4.694
0900A	4	5.959	5.714	7.449	5.906
5042A	7.907	6.676	5.997	6.603	7.552
0483A	3.826	3.827	2.997	6.364	6.779
0105A	2.907	2.976	3	2.871	3.827
2048A	2.913	3.614	3.923	4.826	5.561
0065A	4	3	4.923	4.827	7.642
4844	2.907	1.7	1.927	5.1	4.633
0138A	9.797	5.969	5.782	10.346	12.667
2858A	2.994	1.914	2.991	4.742	3
0585A	3.994	4.827	3	5.985	5.907
0715A	1.999	3.821	3.851	8.774	8.814
0649A	4.82	2.994	3.652	4.309	5.565
4108A	3.994	3.694	4.702	4.463	4.314
0243A	4.953	4.647	3.927	6.654	5.651
0064A	6.895	5.575	7.705	7.53	8.807
0026A	1	1	1	1	1
2425A	5.913	6.306	4.977	5.228	4.614
1926A	5.952	5.688	4.938	8.357	5
1388A	2.994	2.907	3	3.622	3.614
2434A	2.907	2.914	2.978	4.88	5.802
5004A	9.825	9.093	8.146	6.907	7.289
0264A	2	2	2.941	4.614	5.888
0752A	4	5.739	4.837	3.743	4.907
Average	4.213469	4.123125	4.129563	5.5195	5.673313

5.3.4 F_{is}

Levels of inbreeding appear to be highest at Mgambo and Saadani, with the lowest amount at Ikoma. The highest level of inbreeding at the Mara sites was at the Serengeti. Values within sites in the same regions are comparable (*Table 28*).

Table 28. F_{is} across the five populations using 32 loci.

F_{is} per population					
Loci	**Ikoma**	**Ikorongo**	**Serengeti**	**Mgambo**	**Saadani**
1052A	0.069	0.011	0.07	0.143	0.253
1383A	0.876	0.915	0.809	0.964	0.884
2021A	1	1	0.843	0.934	1
0027A	-0.024	1	0.88	0.957	0.823
0702A	0	-0.113	-0.207	-0.018	0.241
2107A	-0.023	-0.05	-0.026	-0.013	-0.098
3840A	0.796	0.947	0.938	0.922	0.872
1333A	-0.026	-0.175	0.382	-0.14	-0.172
1445A	0.825	0.87	0.81	0.954	0.908
0900A	0.859	0.942	0.828	0.7	0.863
5042A	-0.068	0.091	0.043	0.063	0.134
0483A	0.706	0.831	1	0.959	0.913
0105A	0.662	0.776	0.91	0.944	0.866
2048A	-0.215	-0.077	-0.175	0.062	-0.15
0065A	0.817	0.836	0.861	0.914	0.955
4844	-0.031	0	-0.018	-0.002	0.065
0138A	0.689	0.953	0.952	0.933	0.892
2858A	0.873	1	1	0.892	0.944
0585A	0.793	0.906	0.897	0.964	0.916
0715A	1	1	1	0.959	0.908
0649A	0.216	0.183	-0.205	0.193	-0.091
4108A	-0.285	-0.109	0.058	-0.337	-0.011
0243A	0.789	0.889	0.892	0.956	0.949
0064A	0.869	0.948	1	1	0.882
0026A	NA	NA	NA	NA	NA
2425A	-0.03	-0.139	0.053	-0.162	0.075

1926A	0.34	0.21	0.215	0.518	0.534
1388A	1	1	0.809	0.929	0.91
2434A	0.662	1	0.848	0.937	0.852
5004A	-0.046	0.024	0.002	0.042	0.042
0264A	0.709	1	1	0.948	0.909
0752A	0.794	0.901	0.828	0.89	0.95
Average	0.456	0.523	0.542	0.62	0.62

The proportion of randomisation which gave a larger F_{is} than the observed shows which loci are out of Hardy-Weinberg equilibrium (Table 29). This test was based on 3200 randomisations and the indicative adjusted nominal level -5% for the table is 0.00031. 20 loci were significantly out of H-W for at least 1 population. 18 of these were out in at least 4/5 populations. Only 12 loci were not out of H-W.

Table 29. Proportion of randomisation which gave a larger F_{is} than the observed. Red font highlights a loci which is significantly out of H-W equilibrium in a population.

Proportion of randomisation which gave a larger F_{is} than the observed					
Loci	**Ikoma**	**Ikorongo**	**Serengeti**	**Mgambo**	**Saadani**
1052A	0.3678	0.5303	0.2809	0.0947	0.0366
1383A	0.0003	0.0003	0.0003	0.0003	0.0003
2021A	0.0022	0.0006	0.0009	0.0003	0.0003
0027A	1	0.0003	0.0003	0.0003	0.0003
0702A	0.6159	0.8722	1	0.6269	0.0809
2107A	1	0.7259	0.6934	0.6444	0.8475
3840A	0.0003	0.0003	0.0003	0.0003	0.0003
1333A	0.7172	0.9606	0.0225	0.865	0.8825
1445A	0.0003	0.0003	0.0003	0.0003	0.0003
0900A	0.0003	0.0003	0.0003	0.0003	0.0003
5042A	0.8728	0.2059	0.3859	0.3053	0.1244
0483A	0.0006	0.0003	0.0003	0.0003	0.0003
0105A	0.0175	0.0003	0.0003	0.0003	0.0003
2048A	0.9197	0.7953	0.9422	0.3678	0.9109
0065A	0.0003	0.0003	0.0003	0.0003	0.0003

4844	1	1	1	0.5872	0.4041
0138A	0.0003	0.0003	0.0003	0.0003	0.0003
2858A	0.0003	0.0141	0.0003	0.0003	0.0003
0585A	0.0003	0.0003	0.0003	0.0003	0.0003
0715A	0.0241	0.0003	0.0003	0.0003	0.0003
0649A	0.1381	0.1713	1	0.1053	0.9044
4108A	0.9872	0.8294	0.42	1	0.6109
0243A	0.0003	0.0003	0.0003	0.0003	0.0003
0064A	0.0003	0.0003	0.0003	0.0003	0.0003
0026A	NA	NA	NA	NA	NA
2425A	0.6909	0.9291	0.3931	0.9641	0.3369
1926A	0.0113	0.0575	0.0591	0.0003	0.0003
1388A	0.0003	0.0003	0.0003	0.0003	0.0003
2434A	0.0213	0.0003	0.0003	0.0003	0.0003
5004A	0.8031	0.4575	0.5734	0.3897	0.415
0264A	0.0066	0.0003	0.0003	0.0003	0.0003
0752A	0.0003	0.0003	0.0003	0.0003	0.0003
All	0.0003	0.0003	0.0003	0.0003	0.0003
Total loci < P	**12**	**17**	**18**	**20**	**20**

It is possible that the significant deviations from Hardy-Weinberg equilibrium are caused by null alleles. Subsequent analysis is therefore conducted using both the 32 loci data set, and an 11 loci data set which excludes the 20 loci which deviate from H-W and also excludes 0026A was which monomorphic.

5.3.5 LD

For the 32 loci dataset, out of 496 loci pair combinations, 38 (7.6%) of them were significantly in linkage disequilibrium. 29 of 32 loci were in at least one pair which was significantly in linkage disequilibrium. 11 loci were in only one significant pair, 10 loci were in two pairs, two loci were in three pairs, three loci were in five pairs, one loci was in six pairs, one loci was in seven pairs and one loci was in 11 pairs.

For the 11 loci dataset, out of 55 loci pair combinations, two (3.6%) of them were significantly in linkage disequilibrium. Three loci were in at least one pair which was

significantly in linkage disequilibrium. One loci was in two significant pairs, and the other two loci were in only one significant pair.

5.3.6 F_{st}

With 32 loci, the greatest F_{st} value is between Ikorongo and Saadani and the lowest Fst value is between Mgambo and Saadani (Table 30).

Table 30. F_{st} values between the five populations using 32 loci.

	1 (Ikoma)	2 (Ikorongo)	3 (Serengeti)	4 (Mgambo)
2 (Ikorongo)	0.007935			
3 (Serengeti)	0.00713	0.004612		
4 (Mgambo)	0.1221	0.1314	0.1304	
5 (Saadani)	0.1375	0.146	0.1449	0.0006153

With 11 loci, the greatest F_{st} value is between Serengeti and Mgambo, and the lowest is between Ikoma and Serengeti (Table 31).

Table 31. F_{st} values between the five populations using 11 loci.

	1 (Ikoma)	2 (Ikorongo)	3 (Serengeti)	4 (Mgambo)
2 (Ikorongo)	0.003288			
3 (Serengeti)	-0.000844	0.004489		
4 (Mgambo)	0.1336	0.1534	0.157	
5 (Saadanii)	0.1284	0.1472	0.1526	0.0004979

5.3.7 FSTAT between groups

There were no significant differences in allelic richness, observed heterozygosity, gene diversity, F_{is}, F_{st}, relatedness or corrected relatedness between a group containing the three Mara sites and a group containing the two Tanga sites (Table 32) based on 10,000 permutations. Although not significant; allelic richness, observed heterozygosity and gene diversity were all greater in the Tanga sites. Although F_{is} difference was also not significant, a higher F_{is} in the Tanga group could indicate more of a family structuring.

Table 32. Two sided P-values between the Mara populations and the Tanga populations.

Two sided P-values	
Allelic Richness:	0.1739
Ho:	0.4896
Hs:	0.0973
F_{is}:	0.327
F_{st}:	0.9351
Rel:	0.9351
Relc:	0.327

5.3.8 Bottleneck

When the bottleneck software was run with 32 loci, all the populations were significant for a population expansion rather than a genetic bottleneck (Table 33).

Table 33. Bottleneck results for the five populations using 32 loci.

Population	Probability (one tail for H deficit)	Probability (one tail for H excess)
Ikoma	0.00108	0.99900
Ikorongo	0.00002	0.99999
Serengeti	0.00002	0.99998
Mgambo	0.00032	0.99970
Saadani	0.00023	0.99979

This result did not change when using only the 11 loci data set; all were still significant for a population expansion (Table 34).

Table 34. Bottleneck results for the five populations using 11 loci.

Population	Probability (one tail for H deficit)	Probability (one tail for H excess)
Ikoma	0.02686	0.98950
Ikorongo	0.00244	0.99829
Serengeti	0.00464	0.99658
Mgambo	0.00049	0.99976
Saadani	0.00073	0.99951

5.3.9 ML-Relate

Using the 11-loci data set in a GENEPOP format for individual populations, the relationship between flies within populations was examined. It is Mgambo and Saadani which appear to have the most related population, followed by Ikorongo and Serengeti (Table 35). However, none of these differences are significant. There was also no significant differences across populations.

Table 35. Results from ML-Relate using 11 loci.

Within populations	U (%)	HS (%)	FS (%)	PO (%)	Related (%)
Ikorongo	83.90805	12.18391	1.83908	2.068966	16.09195
Ikoma	85.77075	8.300395	1.581028	4.347826	14.22925
Serengeti	81.7734	13.30049	1.231527	3.694581	18.2266
Mgambo	83.14394	12.5	2.840909	1.515152	16.85606
Saadani	86.2069	9.195402	1.149425	3.448276	13.7931

5.3.10 Ne Estimator

All five populations gave Ne with 95% upper confidence intervals of 'Infinite' so the test was not capable of providing reliable results for these populations.

5.3.11 BAPS

Using 32 loci, the three Mara populations were grouped together with no individuals appearing related to the Tanga populations. The Tanga populations appear to be very mixed, with only a few individuals from the same clusters present in each location (Figure 65).

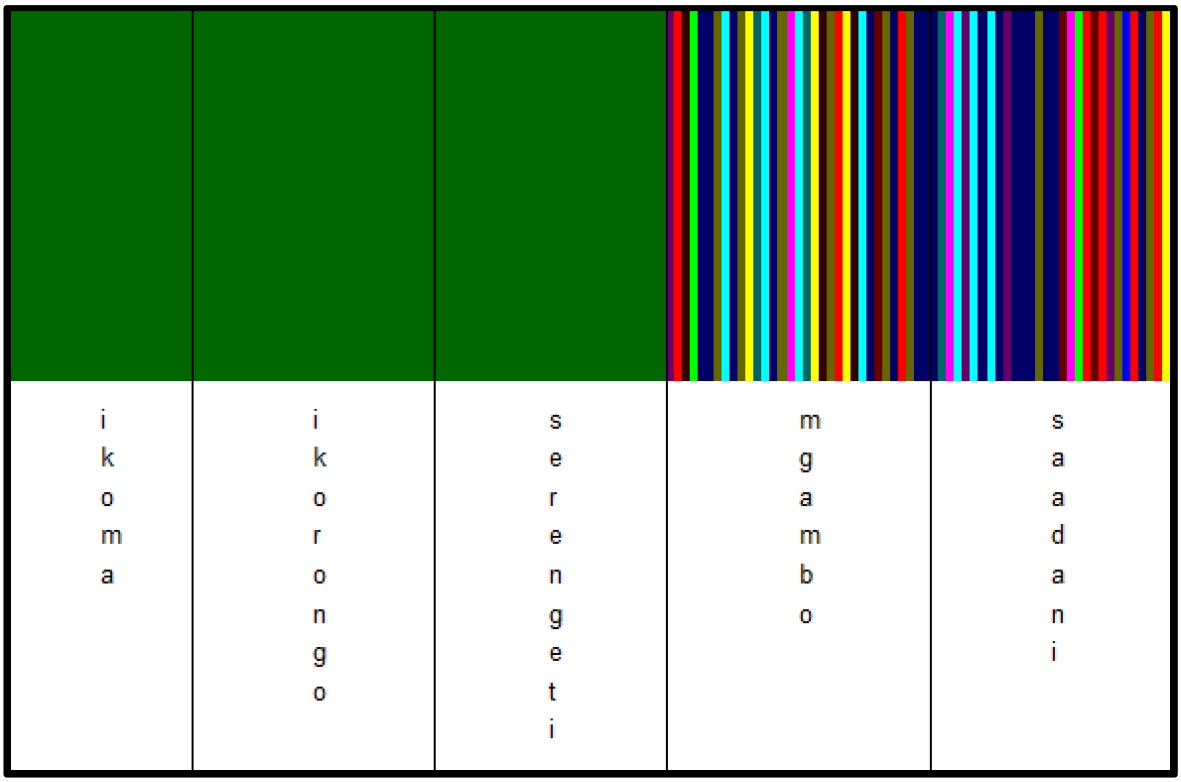

Figure 65. BAPS clustering using the 32 loci dataset.

Using 11 loci, the populations were divided into three clusters. The majority of the Mara populations have all been placed in the same cluster, however there is one individual at Ikoma which appears to be different (Figure 66). The cluster it has been placed in is found at both the Tanga sites, which appear to have a roughly 50:50 split of two clusters. Using BAPS admixture analysis, this showed that there had been no recent gene flow been the three clusters.

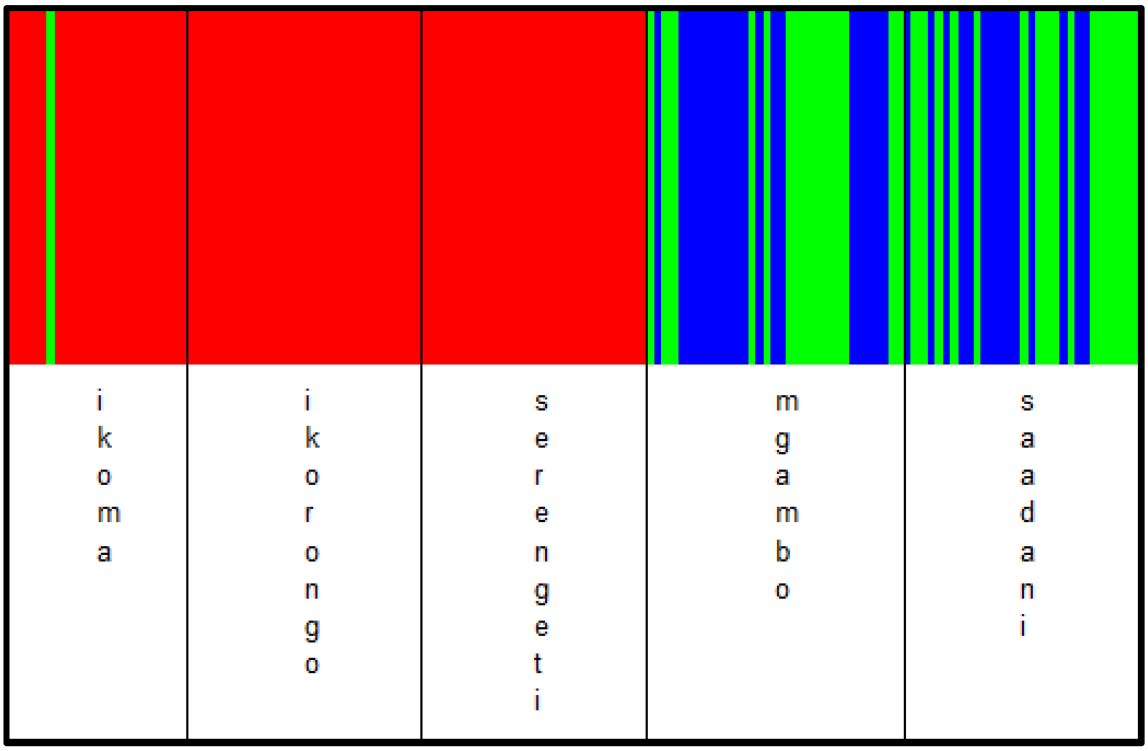

Figure 66. BAPS clustering based on the 11 loci data set.

Using only 11 loci, there is very little structuring which can be determined within geographical locations. To determine if any other loci could be included in the more detailed STRUCTURE analysis FreeNA was used.

5.3.12 FreeNA

FreeNA was first used to quantify the presence of null alleles to determine whether they could be used in STRUCTURE. F_{st} values between populations were compared with null alleles both included and excluded. There were no obvious outliers and the results gave an R^2 value of 0.9934 (Figure 67) for all loci which indicates that the F_{st} values with and without the use of null alleles are highly correlated. This justifies the inclusion of the loci with apparently null alleles to retain extra loci which could give a clearer picture of differentiation.

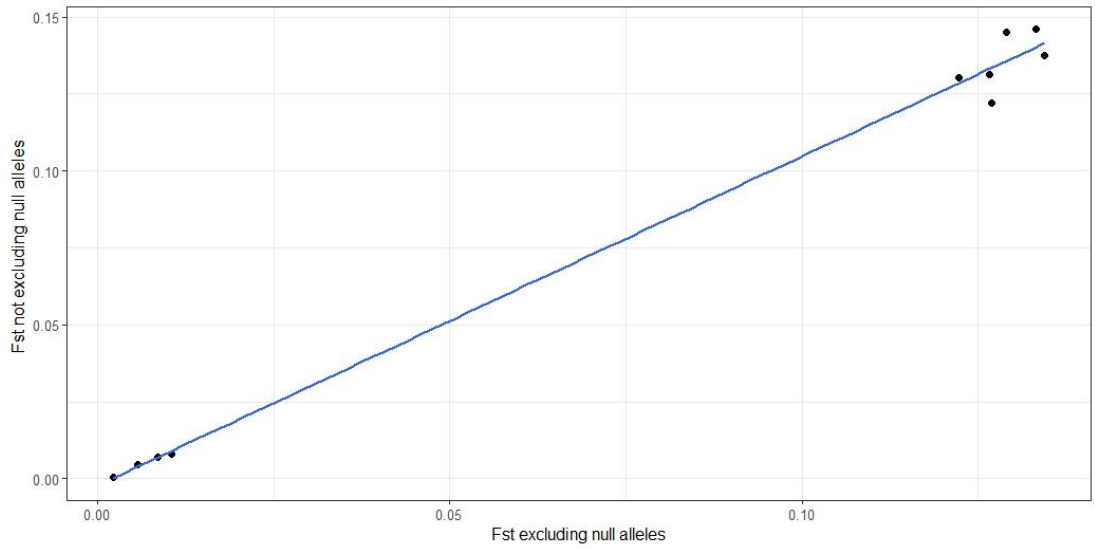

Figure 67. F_{st} values gained using the FreeNA software with and without null alleles.

5.3.13 STRUCTURE

Using 32 loci across the five populations STRUCTURE Harvester suggests that K = 4 using the Evanno method. The CLUMPAK output visualises the results (Figure 68). This shows the three Mara populations as looking very similar, and the two Tanga populations appear to have a similar composition to each other, but they appear more variable within samples than the Mara samples. K=2 has split the samples simply by geographical area, while K=3-5 look like feasible options. The amount of variation between individuals in the Tanga sites within K=3-5 looks indicative of a mixed population.

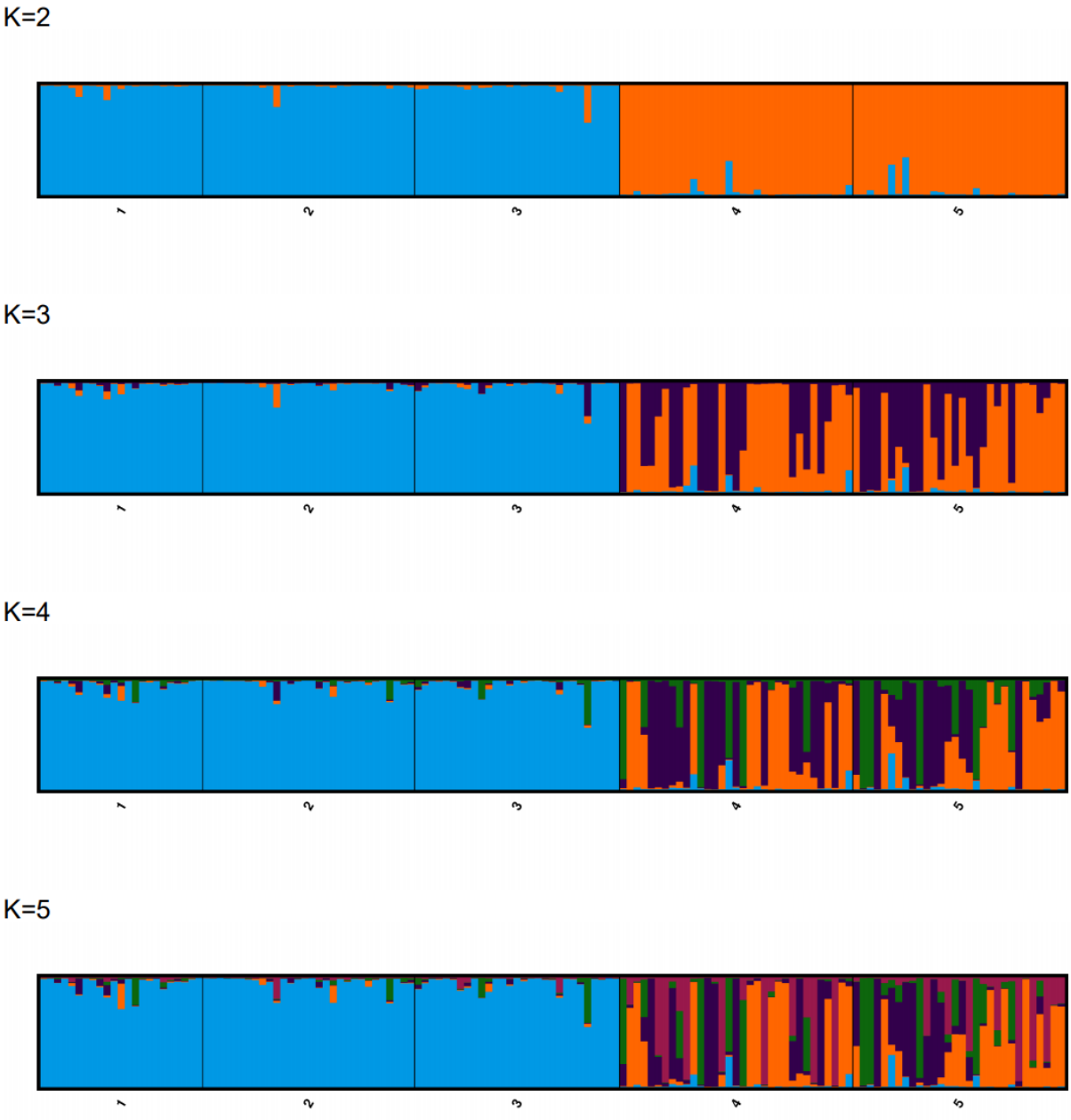

Figure 68. CLUMPAK output from STRUCTURE results from K=2 to K=5 using 32 loci. K indicates number of clusters.

When only 11 loci are used, STRUCTURE harvester suggests K=3. The differentiation shown by the 32 loci BAPS output no longer appears to be present, the major differences are only between the Mara and Tanga groups (Figure 69). K=3 appears to be shuffling noise within the Mara samples and therefore with 11 loci the most appropriate K looks to be 2.

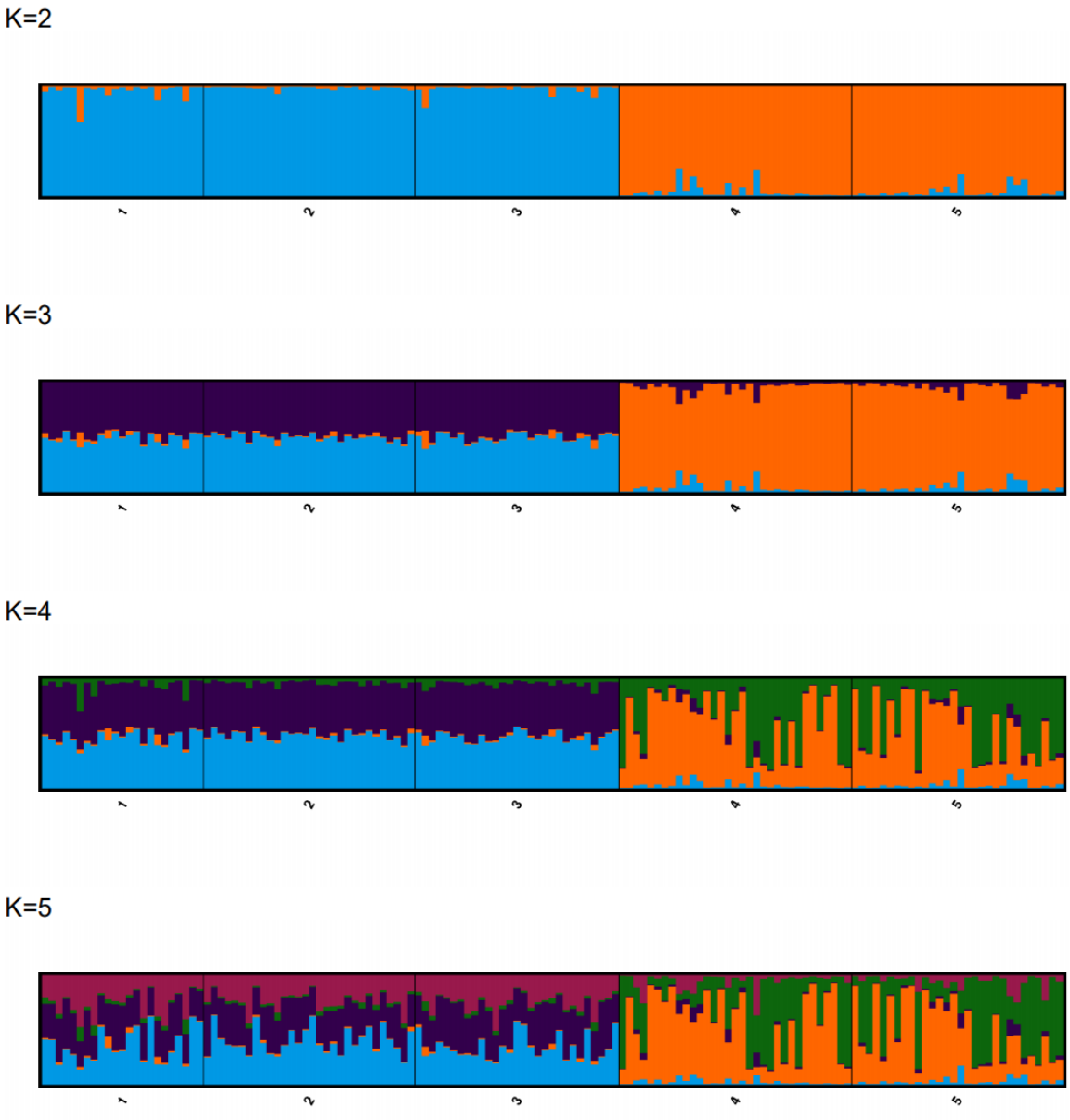

Figure 69. CLUMPAK output from STRUCTURE results from K=2 to K=5 using 11 loci. K indicates number of clusters.

To determine if there is any additional structuring within the two coastal populations, they were run together without the Mara sites (Figure 70). STRUCTURE harvester suggested that K=3. The variation between the two sites looks to be split fairly equally as it did in BAPS.

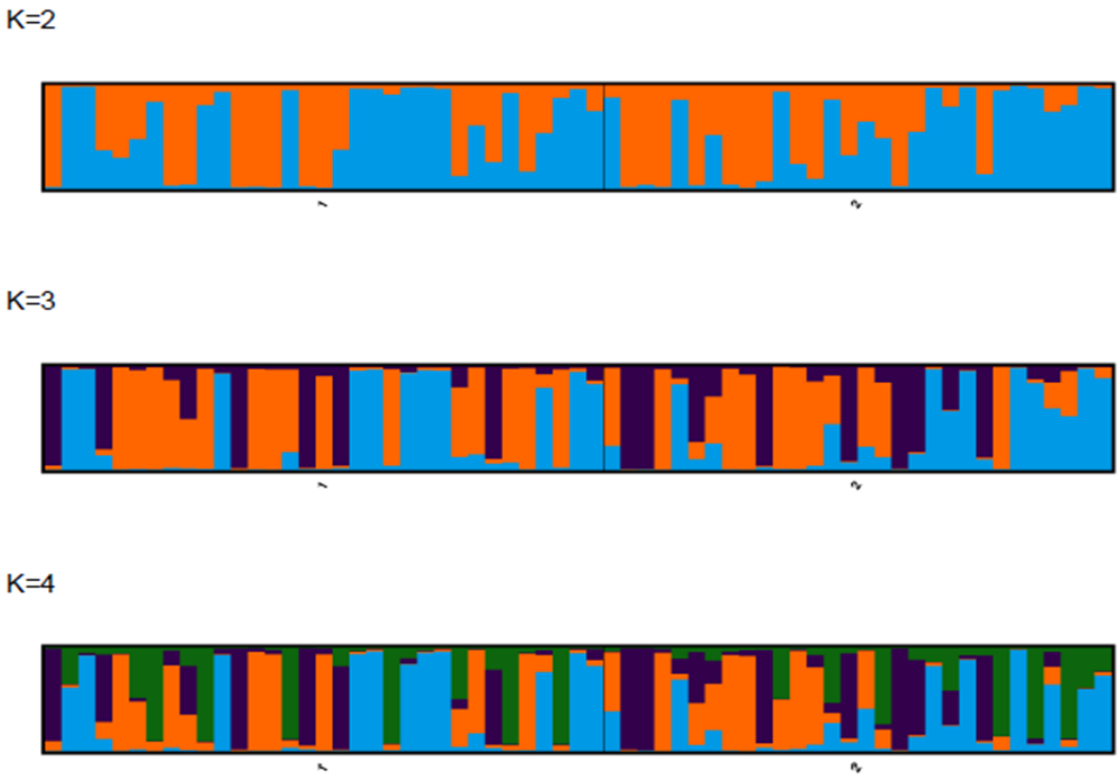

Figure 70. CLUMPAK visualisation of STRUCTURE analysis of the two Tanga populations using 32 loci. K indicates number of clusters.

When the two Tanga populations were run with only 11 loci, STRUCTURE harvester stated that K=2 was the most likely option. However, K=2 looks like it is shuffling noise within the samples (Figure 71).

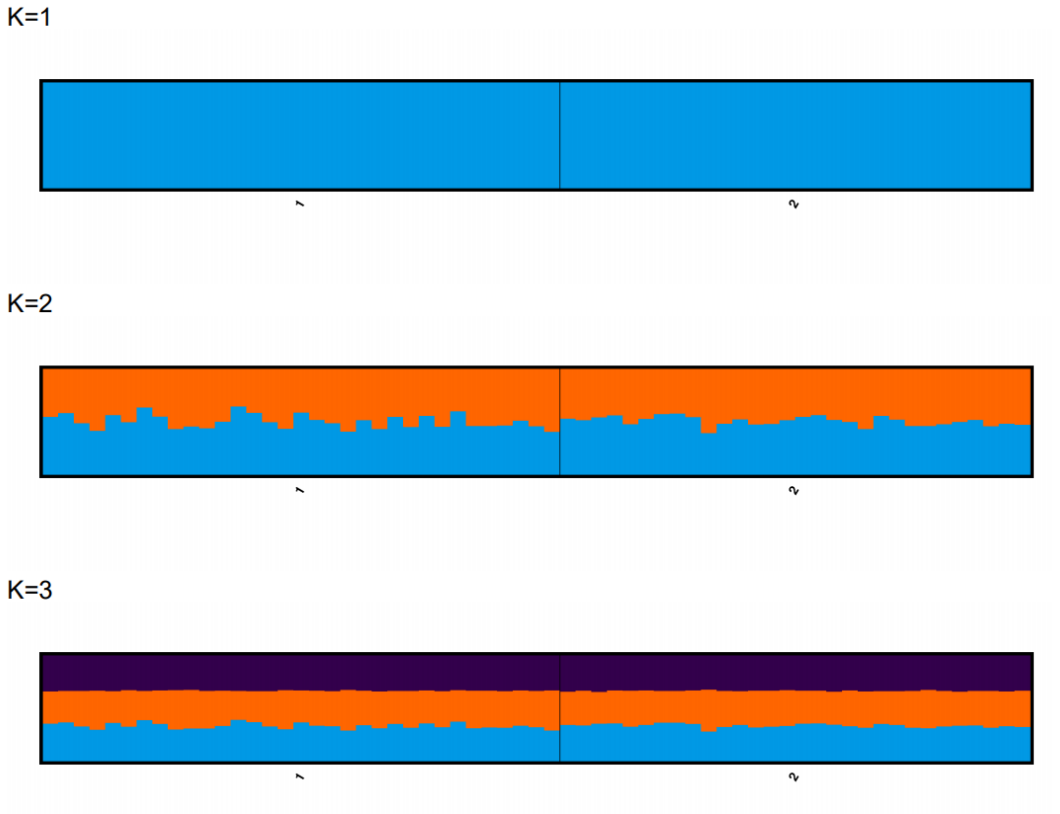

Figure 71. CLUMPAK visualisation of STRUCTURE analysis of the two Tanga populations using 11 loci. K indicates number of clusters.

5.3.14 New clusters

Based on the BAPS and STRUCTURE arrangements, it appears that there may be several populations which deviate from the panmictic model. The individuals were then re-grouped based on their BAPS clusters, with new sample sizes shown in Table 36 and Table 37 . The individual fly at Ikoma in cluster C has been removed from most of the following analysis when split by location due to sample size.

Table 36. New sample sizes based on BAPS cluster.

Population	N
Cluster B (Serengeti)	**81**
Cluster A (Mgambo and Saadani split)	**26**
Cluster C (Mgambo, Saadani +1 Ikoma)	**38**

Table 37. New sample sizes based on BAPS cluster and sampling location.

Population	N
B Ikoma	**22**
B Ikorongo	**30**
B Serengeti	**29**
A Mgambo	**14**
A Saadani	**12**
C Ikoma	**1**
C Mgambo	**19**
C Saadani	**18**

5.3.15 Gene diversity

Gene diversity across the three clusters appears to be evenly distributed with no significant differences between clusters (Table 38). The Serengeti cluster does however have the lowest gene diversity and the two clusters containing the Mgambo and Saadani populations are higher.

Table 38. Gene diversity across the three BAPS clusters and separated by location with 11 loci.

Gene diversity per locus and population										
Loci	**B**	**A**	**C**	**B Ikoma**	**B Ikorongo**	**B Serengeti**	**A Mgambo**	**A Saadani**	**C Mgambo**	**C Saadani**
1052A	0.729	0.508	0.478	0.699	0.707	0.778	0.615	0.375	0.58	0.348
0702A	0.432	0.595	0.569	0.424	0.509	0.355	0.541	0.663	0.627	0.515
2107A	0.316	0.28	0.542	0.133	0.413	0.336	0.206	0.364	0.501	0.603
1333A	0.404	0.38	0.542	0.369	0.454	0.390	0.346	0.424	0.523	0.56
5042A	0.817	0.578	0.783	0.844	0.807	0.821	0.593	0.583	0.806	0.732
2048A	0.467	0.651	0.633	0.500	0.402	0.499	0.640	0.664	0.655	0.614
4844A	0.072	0.505	0.590	0.132	0.033	0.068	0.508	0.515	0.561	0.621
0649A	0.447	0.528	0.427	0.495	0.49	0.372	0.596	0.462	0.436	0.394
4108A	0.565	0.562	0.406	0.589	0.481	0.623	0.626	0.477	0.38	0.454
2425A	0.689	0.668	0.591	0.714	0.673	0.692	0.591	0.739	0.608	0.552
5004A	0.832	0.594	0.764	0.865	0.854	0.76	0.475	0.712	0.787	0.73
Average	0.525	0.532	0.575	0.524	0.529	0.518	0.522	0.543	0.588	0.557

5.3.16 Allelic Richness

Allelic richness was lowest in the first Mgambo/Saadani cluster (A), and highest in the other Mgambo/Saadani cluster (C) (Table 39). When separated by location as well, the group with the lowest allelic richness was for cluster A at Mgambo, and the group with the highest average was for cluster B at Ikoma.

Table 39. Allelic richness across the three BAPS clusters and separated by location with 11 loci.

Allelic Richness per locus and population based on min. sample size of 25 diploid individuals										
Loci	**B**	**A**	**C**	**B Ikoma**	**B Ikorongo**	**B Serengeti**	**A Mgambo**	**A Saadani**	**C Mgambo**	**C Saadani**
1052A	6.638	6.885	6.161	5.586	5.071	6.264	6.103	3.909	5.168	3.45
0702A	3.521	3	3.666	2.9	3.332	2.782	2.96	3	3.554	2
2107A	3.876	4.885	4.646	2.5	3.35	3.455	3.357	2.917	3.567	3.932
1333A	3.622	2	3.646	2.499	3.503	2.757	2	2	2.934	3.222
5042A	6.67	4	7.35	6.613	6.034	5.823	3.828	3	5.876	7.001
2048A	3.671	5	5.429	2.5	2.968	3.479	3.921	5	3.576	4.078
4844A	2.155	4.923	3.96	2.256	1.367	1.619	3.78	3.833	3.408	3.765
0649A	3.878	3	5.572	3.756	2.848	2.993	3	2.996	3.736	4.638
4108A	4.194	4	3.316	3.702	3.215	4.009	3.954	3.913	2.579	2.611
2425A	5.958	3.999	5.312	5.214	5.093	4.582	3	3.996	5.074	3.608
5004A	8.838	6.96	7.198	8.032	7.761	6.615	3.779	5.83	6.066	5.41
Average	4.820	4.423	5.114	4.142	4.049	4.034	3.607	3.672	4.140	3.974

5.3.17 F_{is}

Levels of inbreeding when comparing across clusters shows this is highest in cluster C (Table 40). When broken down by location and cluster this is highest jointly in the Saadani flies from both cluster A and C.

Table 40. F_{is} for each of the three BAPS clusters and separated by location with 11 loci.

F_{is} per population										
Loci	**B**	**A**	**C**	**B Ikoma**	**B Ikorongo**	**B Serengeti**	**A Mgambo**	**A Saadani**	**C Mgambo**	**C Saadani**
1052A	0.052	0.168	0.175	0.09	0.011	0.07	0.188	0.111	0.093	0.362
0702A	-0.114	0.16	0.051	-0.011	-0.113	-0.207	-0.056	0.371	-0.007	0.086
2107A	-0.015	0.176	-0.165	-0.024	-0.05	-0.026	-0.04	0.313	-0.05	-0.29
1333A	0.053	-0.316	-0.117	0.015	-0.175	0.382	-0.238	-0.375	-0.106	-0.09
5042A	0.02	0.101	0.068	-0.077	0.091	0.043	0.222	0	-0.045	0.196
2048A	-0.137	0.078	-0.122	-0.182	-0.077	-0.175	0.219	-0.096	-0.045	-0.176
4844A	-0.027	-0.067	0.083	-0.033	0	-0.018	-0.265	0.191	0.156	-0.041
0649A	0.089	-0.06	0.137	0.265	0.183	-0.205	0.097	-0.262	0.275	0.012
4108A	-0.093	-0.368	-0.038	-0.313	-0.109	0.058	-0.482	-0.222	-0.246	0.144
2425A	-0.04	-0.21	0.066	-0.018	-0.139	0.053	-0.451	-0.015	0.048	0.095
5004A	0.006	0.093	-0.034	-0.051	0.024	0.002	0.098	0.064	-0.071	0.011
All	-0.014	-0.024	0.006	-0.037	-0.025	0.012	-0.058	0.014	-0.001	0.014

The proportion of randomisation which gave a larger F_{is} than the observed for the three BAPS clusters was calculated based on 660 randomisations and the indicative adjusted nominal level -5% is 0.00152. For the populations separated by BAPS and location the calculations were based on 1540 randomisations and the indicative adjusted nominal level -5% is 0.00065. For both data sets, none of the P values were under their respective adjusted nominal levels (Table 41).

Table 41. P values for randomisations which gave larger F_{is} than the observed for the three clusters and populations separated by location.

Proportion of randomisations that gave a larger F_{is} than the observed										
Loci	**B**	**A**	**C**	**B Ikoma**	**B Ikorongo**	**B Serengeti**	**A Mgambo**	**A Saadani**	**C Mgambo**	**C Saadani**
1052A	0.217	0.121	0.079	0.303	0.527	0.290	0.188	0.397	0.342	0.055
0702A	0.932	0.188	0.438	0.653	0.885	1.000	0.690	0.072	0.616	0.560
2107A	0.656	0.218	0.974	1.000	0.746	0.694	1.000	0.263	0.745	0.994
1333A	0.332	1.000	0.867	0.650	0.960	0.029	1.000	1.000	0.790	0.750
5042A	0.382	0.318	0.252	0.879	0.197	0.364	0.212	0.631	0.747	0.086
2048A	0.946	0.362	0.880	0.897	0.779	0.949	0.186	0.801	0.704	0.899
4844A	1.000	0.770	0.336	1.000	1.000	1.000	1.000	0.310	0.281	0.688
0649A	0.211	0.727	0.139	0.090	0.168	1.000	0.430	1.000	0.109	0.632
4108A	0.876	1.000	0.692	0.990	0.825	0.388	1.000	1.000	1.000	0.380
2425A	0.767	0.976	0.306	0.682	0.925	0.409	1.000	0.668	0.468	0.366
5004A	0.523	0.288	0.718	0.846	0.447	0.572	0.417	0.462	0.812	0.585
All	0.733	0.697	0.396	0.823	0.784	0.385	0.816	0.434	0.544	0.392

5.3.18 HW for excluded loci in new clusters

The 20 excluded loci were still significantly out of H-W across all three clusters.

5.3.19 LD

The 32 loci data set was tested after re-grouping and out of 496 tests a total of 59 loci pairs (11.9%) were significantly in linkage disequilibrium. For the 11 loci dataset for the three clusters, out of 55 loci pair combinations, two (3.6%) of them were significantly in linkage disequilibrium. Three loci were in at least one pair which was significantly in linkage disequilibrium. One loci was in two significant pairs, and the other two loci were in only one significant pair. The three loci involved were different to the three loci which were significantly in linkage disequilibrium in the analysis of the original five populations.

5.3.20 F_{st}

5.3.20.1 Three clusters

Between the three BAPS clusters, the greatest F_{st} value was between B and A and the lowest F_{st} value was between A and C (Table 42).

Table 42. F_{st} values between the three clusters with 11 loci.

	B	A
A	**0.1595**	
C	**0.1462**	**0.03**

5.3.20.2 Eight populations

When separated by BAPS cluster and location, the Serengeti samples in cluster B and the Mgambo samples in cluster A had the greatest F_{st} value (Table 43). The lowest F_{st} value was between the individual at Ikoma in cluster C and the other Ikoma flies in cluster B. If the individual at Ikoma is discounted, the next lowest is between flies in cluster B at Ikoma and the Serengeti.

Table 43. F_{st} values between populations separated by cluster and location with 11 loci.

	B Ikoma	B Ikorongo	B Serengeti	A Mgambo	A Saadani	C Ikoma	C Mgambo
B(Ikorongo)	0.00167						
B (Serengeti)	-0.00223	0.00449					
A (Mgambo)	0.154	0.177	0.1811				
A (Saadani)	0.127	0.1505	0.1556	0.00189			
C (Ikoma)	-0.0424	-0.00694	-0.0213	0.150	0.0802		
C (Mgambo)	0.135	0.151	0.155	0.0311	0.0167	0.0496	
C (Saadani)	0.142	0.156	0.162	0.0441	0.0248	0.0646	-0.00192

5.3.21 FSTAT between groups

There were no significant differences in allelic richness, observed heterozygosity, gene diversity, F_{is}, F_{st}, relatedness or corrected relatedness between three groups containing the all A cluster, B cluster and C cluster samples (Table 44) based on 10,000 permutations.

Table 44. Two sided P-values between the three clusters.

Two sided P-values	
Allelic Richness:	**0.0659**
Ho:	**0.2776**
Hs:	**0.0941**
F_{is}:	**0.5828**
F_{st}:	**0.9968**
Rel:	**0.9936**
Relc:	**0.5828**

5.3.22 Ne Estimator

As for the five original populations, all three clusters gave Ne results with 95% upper confidence intervals of 'Infinite' so the test was again not capable of providing reliable results for these populations.

5.3.23 Bottleneck

5.3.23.1 Three clusters

When the BAPS clusters were examined, cluster B and C were both significant for a population expansion (*Table 45*). Cluster A was not significant for either a population expansion or a bottleneck.

Table 45. Bottleneck results for the three clusters using 11 loci.

Population	Probability (one tail for H deficit)	Probability (one tail for H excess)
Cluster B (Serengeti)	0.00171	0.99878
Cluster A (Mgambo & Saadani)	0.12012	0.89697
Cluster C (Mgambo & Saadani)	0.00024	1.00

5.3.23.2 Eight populations

When examined by BAPS clusters and location, the flies from both Mgambo and Saadani in cluster A, and the flies from Mgambo in cluster C were not significant for an expansion or bottleneck (Table 46). The others were all significant for a population expansion.

Table 46. Bottleneck results across populations separated by cluster and location with 11 loci.

Population	Probability (one tail for H deficit)	Probability (one tail for H excess)
B Ikoma	0.03369	0.97314
B Ikorongo	0.00244	0.99829
B Serengeti	0.00464	0.99658
A Mgambo	0.13916	0.87988
A Saadani	0.06152	0.94922
C Mgambo	0.05078	0.95850
C Saadani	0.00610	0.99536

5.3.24 NJT

5.3.24.1 Three clusters

The neighbour joining tree shows the Serengeti cluster much further genetically from clusters A and C than they are to each other (Figure 72). Clusters A and C look almost equally distant from cluster B.

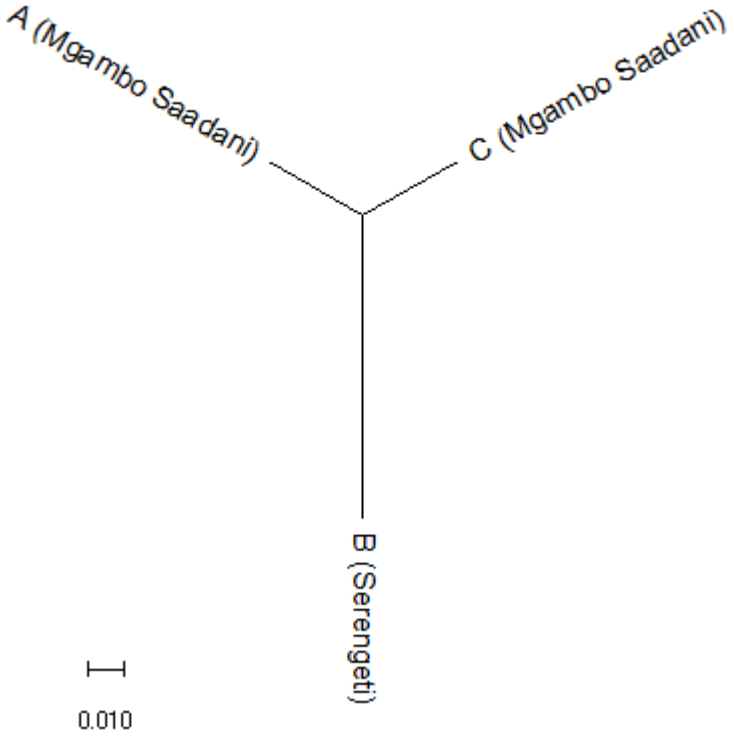

Figure 72. Neighbour joining tree for the three BAPS clusters.

5.3.24.2 Eight populations

The neighbour joining tree for the eight population group shows a similar structure to the one with only the three clusters (Figure 73). The flies from cluster C from Saadani and Mgambo, have been placed genetically closer together than flies which they are geographically closer to from cluster A. The flies from Ikoma in cluster B are genetically closer to those from the Serengeti than those at Ikorongo which is the same as the geographic context. However, the Ikorongo flies are genetically closer to the Serengeti flies than the Ikoma flies which is not the same as the geographic context.

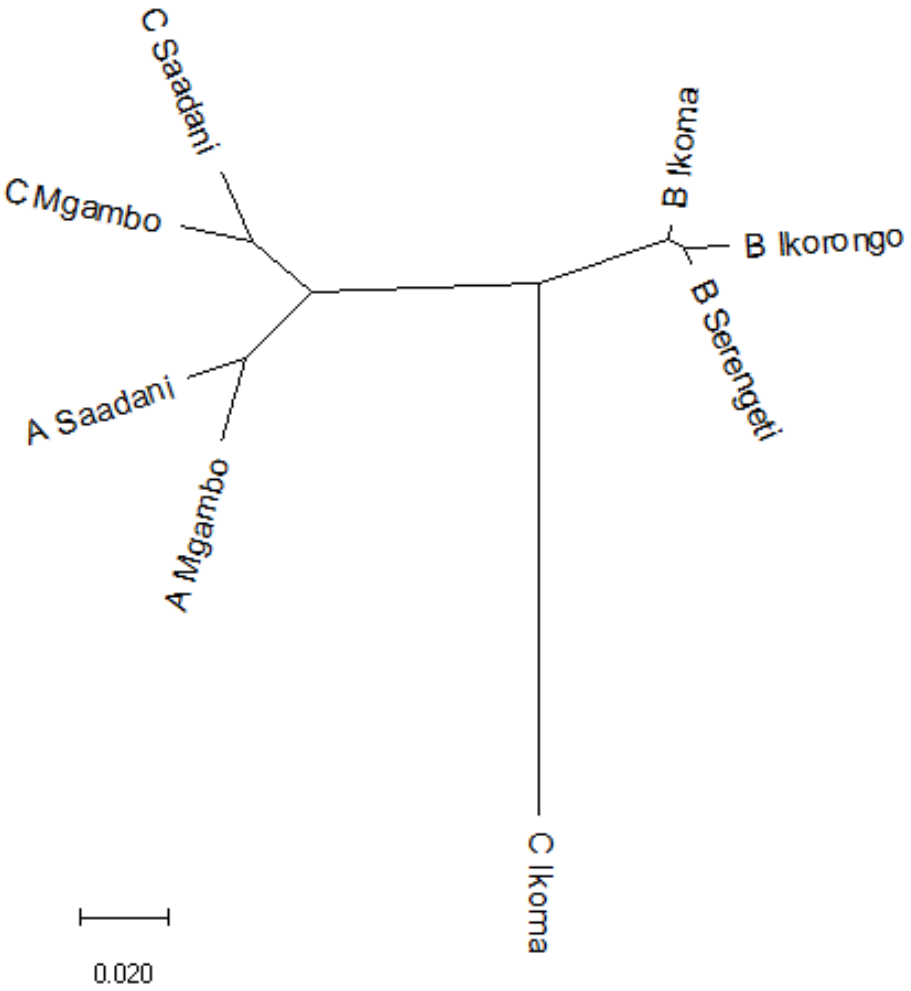

Figure 73. Neighbour Joining Tree for populations separated by BAPS and location.

5.3.25 16S sequencing

Sequencing results were returned for 29 samples, 10 of these were from the Mara cluster (B), 11 from cluster C and 8 from cluster A. All of these samples returning an identification percentage of greater than 98% for *G. pallidipes*. The *G. pallidipes* haplotypes top hits were a selection of ArbaMinch1, Serengeti1, Lambwe7 and TsavoWest2, however many of the second to top hits had the same percentage identity so it was not possible to conclusively define the flies to species haplotype.

5.3.26 AltCOI sequencing

The AltCOI sequences were used to construct a Maximum Likelihood Tree (Figure 74). While it didn't cluster the samples together into the same groups which BAPS did, there are at least no Mara samples in the lower half of the tree which contains the majority of cluster A and C samples. There are however, five samples from A and C which have been clustered into the top half of the tree.

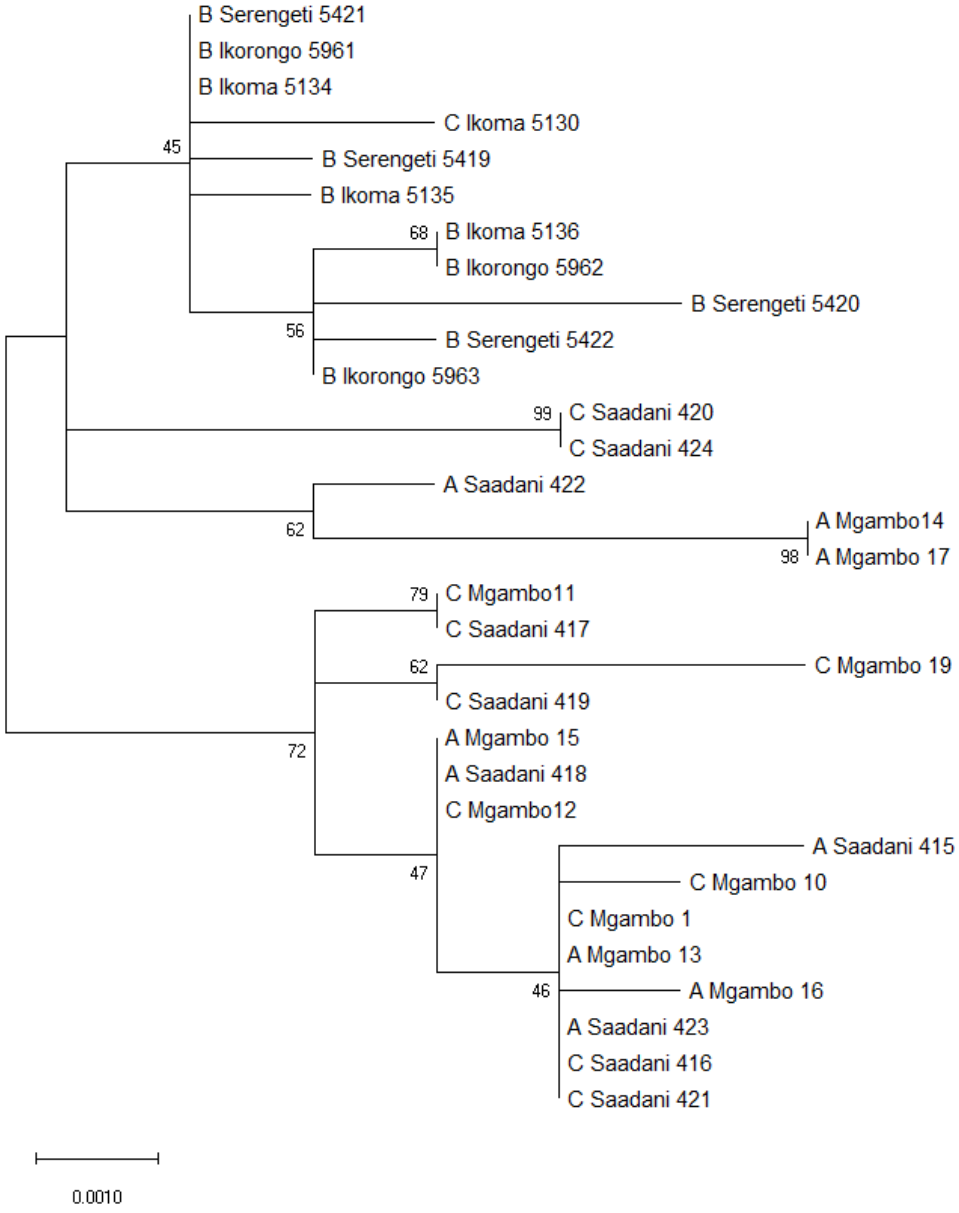

Figure 74. Maximum Likelihood Tree creating using MEGAwith 1000 bootstraps.

5.3.27 WHICHLOCI

When separated into the seven populations (minus the one in cluster C at Ikoma), the 11 loci were only able to correctly assign 64.4% of the samples to their populations. For the three A, B, C clusters – seven loci were able to correctly 95% of the samples and eight loci were able to assign 96% of the samples. The seven loci to assign 95% were 5004A, 4844A, 1052A, 5042A, 2425A, 4108A, 0702A, and the eighth to assign 96% was 2107A. Using all 11 loci could assign 96.3% of samples and this was the maximum percentage which could be correctly assigned.

5.4 Discussion

5.4.1 Main results

In the original 32 loci dataset, the high F_{is} values which were either +1 or very close to +1, indicated a homozygote excess within the populations at many of the loci, which would either be a high measure of population inbreeding, or more likely be due to a technical problem. These technical problems can include null alleles, stuttering, short allele dominance or allele dropouts. Several of the loci with statistically larger F_{is} values were examined in Microchecker (Van Oosterhout *et al.*, 2004) which stated that null alleles were the most likely explanation. This seems a probable explanation as the IAEA strain of flies which the primers were designed from are believed to have originated from the Kenya/Uganda border (Ciosi, Masiga and Turner, 2014) which is closer to the Serengeti sites, and had fewer loci out of H-W equilibrium than at the Tanga sites which are further away. The IAEA colony was established pre-1990 and any genetic differences between these flies may have become more distinct since then, and therefore more likely that mutations may have occurred in areas where these primers were designed for and PCR being unsuccessful for one or both alleles. Eight of the forty original primers had over 25% of their data missing which accounts for 20% of the total initial primers. When the samples were re-grouped and re-tested with the 32 loci data set, the same loci were still deviating from Hardy – Weinberg equilibrium which confirmed that this was a problem with the markers and not a biological effect.

Due perhaps in part to the presence of null alleles, this study was not able to detect significant differences in genetic structuring within geographical regions to address the initial chapter objectives. There is approximately 20 km of unprotected land between the Mgambo and Saadani sites and because of findings in other areas of Tanzania which showed very low fly numbers outside of protected areas and very high levels of control with insecticide, it was hypothesised that the two sites could have become isolated over the last 20 years. Flies from cluster A from Mgambo and Saadani however appear to be very closely related, as do flies from cluster C from Mgambo and Saadani which does not provide evidence that the two sites are isolated from one another. It is unclear whether this is a true biological result, as tsetse are capable of travelling those distances within their lifetimes, or if the markers used were incapable of detecting differences. The bottleneck results were also not entirely expected. When considered just in clusters, only cluster C was significant for a population expansion, cluster A was not. As Saadani was given protected

status from 2005 there are many reasons why the fly populations may have increased in number. While Mgambo may not have the same situation, both A and C clusters were made up of relatively equal numbers of flies from Mgambo (54% in cluster A, 50% in cluster C) and Saadani (46% in cluster A, 47% in cluster C) therefore it is unlikely that flies from Mgambo may have diluted an expansion effect, especially when they are so closely related anyway. When examined by cluster and location, it was only the flies in cluster C at Saadani which were shown to be undergoing a population expansion. Although the flies from cluster A at Saadani and cluster C from Mgambo appeared more likely to have undergone an expansion than a bottleneck, it was still not a significant result.

There were similar levels of differentiation between the Tanga populations as there was between the Mara populations which are definitely connected by protected habitat. The Mara populations all appear to be very closely related. All Mara samples from cluster B were significant for a population expansion. It was hypothesised that in border areas such as Ikorongo, it was likely that due to the high levels of reported control present that a bottleneck may be detected. It was thought this would be less likely at the Serengeti site because the traps were set deeper into the park and should not be coming into contact with insecticide. In fact, it was Ikorongo which had the highest level of significance for an expansion effect. As the data has shown the Mara sites are highly related, so it is likely that flies from Ikoma or the Serengeti are moving into and breeding at the Ikorongo site, and so perhaps an influx of flies from Ikoma and Serengeti are replacing those at Ikorongo which are dying from control, and it is appearing as though the population is growing in diversity. Although not significant, the flies in cluster B at Ikorongo did have the highest level of gene diversity of the three Mara sites.

5.4.2 Control implications

These results have several implications for control. If it is biologically correct, the Mara populations are all highly related. This would render any tsetse control operation targeted within the game reserves useless, as it would be constantly reinvaded from surrounding areas. This highlights the importance of maintaining a barrier around the outside the park as a more effective measure which is currently being done by livestock farmers treating their cattle with insecticide.

The next implication for control relates to the use of SIT. While this has been useful in certain scenarios such as on Zanzibar (Vreysen *et al.*, 2000), there are two reasons why this would not be useful in this area. The first is due to the reinvasion risk as highlighted above, the second is due to the presence of cryptic species. As shown at Mgambo and Saadani,

there are two groups of *G. pallidipes* flies living side by side and not mating. Manangwa *et al.*, (2019) also found two distinct lineages at their Serengeti sites. If a SIT operation were to be planned in Tanzania, it would have to be planned very specifically to the cryptic species present. With the example of Saadani, the operation would be complicated by needing to release two different sets of flies just for *G. pallidipes*, instantly doubling the number of flies needed which may be less acceptable to local people or tourists, especially as male flies also bite, unlike in mosquitoes where only the female bites.

There is also the possibility that the different groups may have slightly different behaviours. Of all the flies collected from Mgambo and Saadani, 59.4% of them belonged to cluster C, and 40.6% of them belonged to cluster A. It could be that they have a smaller population, which would raise questions as to whether they have different habitat suitability. It could however be that they have slightly different host seeking behaviours and Nzi traps may be less effective at catching cluster A flies which would again have implications for control. For example, *G. pallidipes* in Somalia are not caught in significantly greater numbers when chemical attractants are used (Torr, Parker and Leigh-Browne, 1989), however for *G. pallidipes* in Zimbabwe, chemical attractants increased catches by up to 60 times (Vale and Hall, 1985).

5.4.3 Comparisons with Manangwa *et al.*, (2019) paper

Cryptic species are not uncommon in *Glossina* sp.. Dyer *et al.*, (2011) and Echodu *et al.*, (2013) report this for *G. fuscipes*, and also for *G. palpalis palpalis* (Dyer *et al.*, 2008), as does Gooding, Solano and Ravel, (2005). However, only Manangwa *et al.*, (2019) has so far shown this phenomenon in savannah flies.

There are many similarities between this study and the Manangwa *et al.*, (2019) study. Four of their ten loci (40%) failed in over a third of their samples so were mostly excluded from analysis. In this study 20% of the loci were completely excluded after amplification failures. Microchecker suggested that they also had many null alleles, however this did not correlate with their F_{is} values. All of their loci were involved in at least two significant LD pairs, and in this study 91% of loci were involved in at least one significant pair. They also managed to partition their main clades with confidence and was robust with the available loci, but found their results less reliable when within clades, which reflects the results here with WHICHLOCI.

The difference in this study is the linkage disequilibrium results post re-grouping. They found that after re-grouping, none of their locus pairs were in significant LD, whereas in this

study with the 32 loci, the number of locus pairs in significant LD increased. The numbers of significant LD pairs in the 11 loci dataset however stayed constant, highlighting that in this case the LD was related to markers, whereas in the Manangwa *et al.*, (2019) paper this may have been a biological effect. On the map provided in their paper, their Serengeti, Robanda and Ikorongo North sites are geographically closest to the Serengeti, Ikoma and Ikorongo sites respectively in this study. There, they found the majority of flies in one cluster, and two flies from a separate cluster out of 84 flies. The flies from the smaller cluster were found in the Serengeti site and the Ikorongo North site. In this study, one fly was found to belong to another cluster out of 82 flies at the Ikoma site, therefore the proportion of flies belonging to each cluster are comparable. They did however find 7 out of 22 flies in the other smaller cluster at Ikorongo South which is in between Ikorongo and Ikoma.

5.4.4. Limitations and solutions

The first limit of the study was that only males were used. Sex-biased dispersal is common in other species such as *G. fuscipes* (Hyseni *et al.*, 2012), while Ouma *et al.*, (2011) also found a sex-biased dispersal of *G.pallidipes* and concluded that males were dispersing further than females, however, small male samples were used in that study and for the purposes of this study relative comparisons can still be made. Conversely, ecological studies of tsetse using Mark-Release-Recapture methods tend to show that females move more than males (Vale, Hursey, Hargove, *et al.*, 1984).

The second limitation is the ability of the loci to distinguish clusters. While this was possible for 96% of the samples to be grouped into the A, B and C clusters, only 64% of samples were clustered correctly when location was included. Therefore, any analysis broken down by cluster and location must be treated with caution. The 16S mitochondrial sequencing was also not able to distinguish further than *G. pallidipes*, which could be due to the region not being as variable as others and the product being only ~ 300 bp long. The COI primers from (Dyer *et al.*, 2008) were used in an earlier study for *G. pallidipes* (Lea, unpublished) and did not amplify products well. The AltCOI primers (Cunningham, unpublished) were used to see if the same microsatellite clustering could be detected using mitochondrial primers but this was not possible.

To enhance the resolution of the genetic structuring requires more functioning markers in the full data set. For future studies it would be beneficial therefore to sequence several of the homozygotes alongside heterozygotes and identify the mutations causing null alleles. There were also several loci which were common for non-stepwise mutations which was assumed to be scoring errors and were rounded up or down to the most common nearest

allele. However, it may be better to sequence these as well and examine the flanking regions for mutations.

5.5 Conclusion

Based on previous literature where some studies did not detect genetic structuring with less than 10 microsatellite markers, in this study what should have been a comprehensive set of 40 markers was designed so that if structuring was present it could be detected. Unfortunately, due to what is expected to be the presence of many null alleles, several of the loci had to be completely excluded and many viewed as potentially unreliable.

After excluding these unreliable markers, with the resolution available it appears as though, with the exception of one individual, the three Mara populations are one large panmictic population with very little genetic distance between them. It also appears as though there are no genetic bottlenecks at Ikorongo or Ikoma where there was predicted to be, but instead population expansions. Where expansions were predicted at Saadani, this only seems to be partly true.

Despite not being able to address the initial objectives with certainty, this study highlighted an unexpected structuring of cryptic species in Tanga Region which could have serious implications for control efforts. The similarities with the Manangwa *et al.*, (2019) paper confirms that these results are likely to be true, and this study provides further evidence of *G. pallidipes* cryptic species in Tanzania.

Chapter 6: Discussion

While the numbers of HAT cases have been declining in Tanzania from 400/year in 1995 to 3/year in 2017 (WHO, no date a), trypanosomiasis is a continuing threat to both animals and people in the country and across sub-Saharan Africa. The number of cases in humans is likely under-reported in local populations (Simarro *et al.*, 2010), and even a small outbreak in visiting tourists could have a disproportionately large impact on the tourist economy. As the areas affected by human and animal trypanosomiasis are often rural, resource-poor areas, it is important that cost-effective, well-accepted methods of vector control are advocated and utilised by national and local governments.

The first aim of this study was to determine the ability of environmental variables to predict the distribution and abundance of tsetse over small scales which would assist in the design, monitoring and implementation of tsetse control operations, both in areas where data has been collected but also discovering the potential to expand predictions to new areas. A second and closely-related aim was to quantify the extent and levels of control which may have been ongoing, conducted by local people living in affected areas, and, finally, to use molecular methods to examine tsetse populations to see if control could be guided through population genetics. The results of this study have mostly addressed these aims, with sometimes unexpected findings which may have further implications.

6.1 What is the potential for using fine-scale modelling to predict tsetse abundance?

Over large (>1 km) and small (<500 m) scales, numbers of *G. pallidipes* correlated well with remotely-sensed variables such as Band 7, NDVI, Land Surface Temperature and percentage tree cover. Generally, catches were negatively correlated with Band 7 values and distance from a river, positively correlated with NDVI values and percentage tree cover. Many of the environmental variables are related to changes in riverine habitat, however highly vegetated areas without rivers present are also likely to have relatively large numbers of tsetse. *G. swynnertoni* was however spread more homogenously across habitats, although appears to favour more open areas as catches decreased with increasing tree cover. These findings suggest that analysis of satellite data can provide a robust means of identifying monitoring sites for pre- and post-intervention surveys of tsetse populations. Generally, selection of sites depended on an ad hoc selection of sites that are relatively convenient to visit and, on the basis of experience, expected to produce large numbers of tsetse. The findings in this study suggest that a more rational approach could be achieved by selecting

potential monitoring sites which have an optimal combination of, say, Band 7, NDVI, LST and proximity to a river.

A predictive model of tsetse abundance developed from empirical data produced in the Serengeti district, was tested in Tanga region, an area which differed in climate, vegetation and farming systems from Mara region where the model was first developed. Despite these differences, the model performed well at predicting relatively high and low numbers of tsetse within protected areas. The areas where the model did not perform as well was in unprotected areas where tsetse were predicted to be abundant but none were found. The apparent poorer predictive value of the model outside conservation areas may reflect impact of tsetse control (see below). This illustrates the utility of the findings present in this thesis: identification of sites where tsetse are predicted to be abundant will help in the design and implementation of surveys to assess disease risk and the impact of planned or autonomous control of tsetse.

While the Serengeti-based model appears to translate well to protected areas in Tanga region, there are concerns with using a linear model for examining the relationships with environmental variables. Whilst it appears in these study sites the relationships are generally linear, there are likely to be other areas where the values for the environmental variables are greater or less than those examined in this study. Extrapolating past these points means it is not possible to know with certainty how those relationships may change outside of the known range.

There is still however much potential for using modelling to predict relative tsetse abundance at fine scales inside protected areas. In the reality of a control programme, it is likely that a predictive map with such a fine resolution of 100 m would not be of greater use than one with a resolution of 500 m. Control would likely be spread widely across an area and even if a target had not been placed within 100 m of where the highest relative predictions are and was instead 500 m away, it is still likely to impact on flies within that population due to their high rates of movement per day (up to 1 km). It may be more in the context of a monitoring project that a predictive map with a 100 m resolution may be more useful, to ensure that traps have been placed in an area most likely to contain the greatest abundance of flies, which could then be compared with the greatest abundance found in another area and be used to decide on a larger scale where to focus potential control resources in the future.

To develop the model further, it is important to assess its performance in areas where pastoralism predominates. This study has been carried out in areas where either mixed crop-livestock farming (Serengeti district) or zero-grazing (Pangani district) were the dominant forms of livestock production. In these areas, the model predicted flies would be present but, in many cases, appeared to be absent, likely due to the presence of ITC. However, the impact of ITC may be less in pastoralist areas. Pastoralist groups are often very mobile which could contribute to an uneven distribution of control, if as expected they are also treating with insecticide, which could cause a reduced or less constant mortality pressure on tsetse populations (Torr and Vale, 2011).

A robust model of tsetse abundance could then be applied to northern Tanzania to assist the Tanzanian government identify areas most at risk of tsetse. Such information could assist surveys to identify areas where the burden of animal trypanosomiasis is greater and where interventions might be required. For instance, the work of Mbassa *et al.*, (2017), which analysed the potential impact of dip construction and acaricide subsidy on tick-borne disease, could be enhanced by identifying areas where animal trypanosomiasis is likely to be important.

6.2 Is the treatment of cattle with pyrethroids the most cost-effective method of tsetse control in Tanzania?

Questionnaires of livestock keepers and analyses of cattle hair indicated that cattle in Serengeti district are being treated regularly with alphacypermethrin. The GCMS-determined amounts of insecticide appears slightly lower than what was reported, however the percentage of cattle found with insecticide present are likely to be impacting on tsetse populations. The present findings are the first example of chemical analyses of cattle hair being used to assess the proportion of livestock keepers are treating their cattle with pyrethroids. Previous studies have relied solely on qualitative methods such as questionnaires and focus group discussions (Bardosh, Waiswa and Welburn, 2013). The findings are particularly important because they showed that livestock keepers are using a formulation effective against tsetse and that this mass treatment was not part of a concerted and planned tsetse control programme. All previous examples of insecticide-treated cattle being used to control tsetse were part of government- and donor-funded tsetse control programmes (Hargrove *et al.*, 2000; Bardosh, Waiswa and Welburn, 2013).

Now that we know that farmers are using insecticides frequently, the next step is to determine the drivers behind this use. An important focus would be consideration of the economic costs and benefits to livestock keepers, and how these affect decisions about treating their animals with insecticide. In addition to assessing the economic rationale underpinning decisions about treatment of animals, there is also a need to analyse the national supply chain for veterinary insecticides and drugs, how farmers receive and understand information about options for controlling tsetse and trypanosomiasis, and how these factors have changed over the last two decades. Such analyses would help determine if government policies have contributed directly to an increase in insecticide-based control and could give insights on how to improve access to insecticides and therefore further increase uptake of control. This knowledge could be used to help encourage and guide other countries to implement similar policies.

From visiting several agrovet shops in the local villages, it appeared as though suppliers and distributers of veterinary medicine supplies are reaching rural areas and with products effective against both tsetse and ticks. The government is leading with a clear strategy and policies on vector control which appear to have been disseminated effectively. The local livestock officers and veterinarians are delivering this information to livestock owners who are implementing control seemingly effectively.

Work carried out in Serengeti district should be repeated across a range of livestock production systems (e.g., pastoralist, zero-grazing, mixed crop-livestock) and in areas where tsetse are abundant or sparse. In addition to carrying out local studies of insecticide use, obtaining data on sales of veterinary insecticides at local and national levels would provide further information on the scale and extent of pyrethroid use. Together, these studies would reveal whether treatment of cattle with pyrethroids is occurring across Tanzania.

It seems as though this method of tsetse control has been readily accepted by livestock keepers. It is also likely to be favourable for the government, as the subsidy applied at dip tanks is likely to be the most-effective method of disease control compared to other methods such as deploying targets or aerial spraying and is therefore likely to maintained.

The impacts of this frequent use of insecticide may have on tick resistance should however not be ignored. Restricted application is one method which would benefit an integrated control approach for both tick and tsetse by helping to maintain endemic stability (Eisler *et al.*, 2003). However, all livestock keepers spoken to were applying insecticide to the whole

body of the animal. This is a point which could be reinforced by livestock officers when speaking to livestock keepers, with reminders that it would also be financially beneficial.

While there are still many steps between now and the IQK being field-operational, it could be a helpful tool in the future to help with the challenge of tick resistance, as it could be used by livestock officers to make sure that cattle are not treated unnecessarily if insecticide is still present. Although the IQK is not yet suitable for field use, it is very close to becoming an acceptable alternative to HPLC and GCMS in a laboratory setting, which would reduce the cost of quantifying insecticide for monitoring projects. More generally, the IQK could help in monitoring use of insecticides where pyrethroids are controlled to either prevent the development of resistance (Vudriko *et al.*, 2016) or protect the environment (Moore and Waring, 2001).

6.3 What are the implications of finding cryptic species of tsetse in Tanzania?

The main finding of this chapter was the discovery of what appears to be cryptic species of *G. pallidipes* in the Tanga sites. Cryptic species have been found before in several tsetse species (Gooding, Solano and Ravel, 2005; Dyer *et al.*, 2008, 2011; Echodu *et al.*, 2013) and have also very recently been found in the Serengeti area (Manangwa *et al.*, 2019) and those results have been backed up by these findings. A comprehensive set of 40 microsatellite markers was designed, however, due to the presence of null alleles many of the results from these markers were not useful for the intended purpose. In particular, the initial hypotheses of potential genetic bottlenecks due to control at the Serengeti sites could not be proved, and hypotheses of population expansions at Saadani were only seen in one of the two cryptic species found in the area.

The implications for future work are significant as there may be differences in behaviours of each cryptic species which may impact on disease transmission dynamics, their ecologies, and would make SIT programmes particularly difficult. If a simple diagnostic was found, it would be easy to examine differences in the behaviours of wild flies. Responses to certain stimuli such as odour or colour could then be tested across the different cryptic species to examine any differences. Field experiments could be designed using different trapping methods to see if specific factors were more important to each cryptic species, potentially highlighted by the behavioural phenotypic differences in attraction to odours by *G. pallidipes* in Zimbabwe (Vale and Hall, 1985) and Somalia (Torr, Parker and Leigh-Browne, 1989). In this study, only stationary traps were used, however both stationary and mobile

traps could be used, and these could be used in combination with and without different odours. Flies caught in these different ways could then be examined by a simple PCR to determine which cluster they belonged to. Understanding these differences will allow control methods to be tailored to the target species.

The modelling in this thesis and in Lord *et al.*, (2018) assumed that all *G. pallidipes* flies are similar in their behaviour as abundance values were previously gained from all trap catches. It may however be that some cryptic species may be slightly better adapted to living in different habitat types than others and that predictive maps would show higher relative abundances in different areas for each. In addition, responses to traps may vary between cryptic species. For disease transmission modelling, differences in say the responses of cryptic species to different hosts may affect the overall transmission of trypanosomes and ultimately the epidemiology of trypanosomiasis. There is clearly a need to investigate the ecology and behaviour of cryptic species across the entire range of *G. pallidipes*, from Zimbabwe to Somalia.

6.4 Final conclusion

The overarching aim of the studies presented in this thesis was to examine the population ecology of tsetse at the interface of conservation and farming areas in northern Tanzania to assist in the design, implementation and monitoring of interventions to control tsetse. In particular, the studies aimed to quantify the distribution and abundance of tsetse over fine scales (<1 km). It was found that many of the relationships between abundance with environmental variables (Band 7, NDVI, LST) over large scales, are also evident at fine scales within conservation areas where natural vegetation is preserved. There is therefore a great potential for predicting *G. pallidipes* abundances across Tanzania using environmental variables in wilderness areas. In farming areas however, where vegetation is degraded, natural hosts are scarce or absent and livestock may be treated with pyrethroids, the ability to predict even the presence of tsetse using remote sensing is diminished. But even here, use of geostatistical models to identify areas where tsetse are predicted to be abundant will assist in monitoring the impact of vector control programmes and policies on trypanosomiasis. Future control programmes will also have to consider that in areas where *G. pallidipes* is present, it seems that two cryptic species exist, which may also have further implications for the sterile insect technique and, possibly, odour-baited targets.

The findings presented in this thesis also provide the first exciting indications that livestock keepers in Tanzania, at district, regional and national levels, have been successfully controlling trypanosomiasis, both directly and indirectly through vector control facilitated by national policies to promote use of pyrethroids. Further work is required to confirm the extent and impact of these farmer-led interventions, and the development of a cost-effective, simple and rapid system to quantify pyrethroid will play an important part in this. Lessons learned about how farmers can be enabled to control tsetse in Tanzania will provide important lessons for all countries affected by African trypanosomiasis.

References

Adams, E. R. *et al.* (2006) 'Trypanosome identification in wild tsetse populations in Tanzania using generic primers to amplify the ribosomal RNA ITS-1 region', 100, pp. 103–109. doi: 10.1016/j.actatropica.2006.10.002.

Adams, E. R. *et al.* (2008) 'The identification, diversity and prevalence of trypanosomes in field caught tsetse in Tanzania using ITS-1 primers and fluorescent fragment length barcoding', *Infection, Genetics and Evolution*, 8(4), pp. 439–444. doi: 10.1016/j.meegid.2007.07.013.

Agriculture, M. of (2012a) 'The United Republic of Tanzania NATIONAL SAMPLE CENSUS OF AGRICULTURE 2007 / 2008: REGIONAL REPORT : – MARA REGION', *Report to Tanzania Govt*, V(July).

Agriculture, M. of (2012b) 'The United Republic of Tanzania NATIONAL SAMPLE CENSUS OF AGRICULTURE 2007 / 2008: REGIONAL REPORT : – TANGA REGION', *Report to Tanzania Govt*, V.

Aksoy, S. *et al.* (2013) 'Glossina fuscipes populations provide insights for human African trypanosomiasis transmission in Uganda', *Trends in Parasitology*. Elsevier Ltd, 29(8), pp. 394–406. doi: 10.1016/j.pt.2013.06.005.

Alderton, S. *et al.* (2018) 'An agent-based model of tsetse fly response to seasonal climatic drivers : Assessing the impact on sleeping sickness transmission rates', *PLoS neglected tropical diseases*, 12(2), pp. 1–29.

Allan, F. (2019) *East Coast Fever and vaccination at the livestock/wildlife interface*.

Altschul, S. F. *et al.* (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Auty, H. K. *et al.* (2012) 'Using molecular data for epidemiological inference: assessing the prevalence of Trypanosoma brucei rhodesiense in tsetse in Serengeti, Tanzania.', *PLoS neglected tropical diseases*, 6(1), p. e1501. doi: 10.1371/journal.pntd.0001501.

Bacchi, C. J. (2009) 'Chemotherapy of Human African Trypanosomiasis', *Interdisciplinary Perspectives on Infectious Diseases*, 2009, pp. 1–5. doi: 10.1155/2009/195040.

Baldacchino, F. *et al.* (2013) ' Transmission of pathogens by Stomoxys flies (Diptera, Muscidae): a review ', *Parasite*, 20, p. 26. doi: 10.1051/parasite/2013026.

Banks, M. A., Eichert, W. and Olsen, J. B. (2003) 'Which genetic loci have greater population assignment power?', *Bioinformatics*, 19(11), pp. 1436–1438. doi: 10.1093/bioinformatics/btg172.

Barclay, H. J. and Hargrove, J. W. (2005) 'Probability models to facilitate a declaration of pest-free status, with special reference to tsetse (Diptera: Glossinidae)', *Bulletin of Entomological Research*, 95(1), pp. 1–11. doi: 10.1079/ber2004331.

Bardosh, K., Waiswa, C. and Welburn, S. C. (2013) 'Conflict of interest: Use of pyrethroids and amidines against tsetse and ticks in zoonotic sleeping sickness endemic areas of Uganda', *Parasites and Vectors*, 6(1), pp. 1–15. doi: 10.1186/1756-3305-6-204.

Barrett, M. P. *et al.* (2007) 'Human African trypanosomiasis: Pharmacological re-engagement with a neglected disease', *British Journal of Pharmacology*, 152(8), pp. 1155–1171. doi: 10.1038/sj.bjp.0707354.

Barsi, J. A. *et al.* (2014) 'The Spectral Response of the Landsat-8 Operational Land Imager', *Remote sensing*, 6, pp. 10232–10251. doi: 10.3390/rs61010232.

Bauer, B. *et al.* (1995) 'Successful application of deltamethrin pour on to cattle in a campaign against tsetse flies (Glossina spp.) in the pastoral zone of Samorogouan, Burkina Faso', *Tropical Medicine and Parasitology*, 46(3), pp. 183–189.

Bauer, B. *et al.* (1999) 'Improvement of Cattle Productivity through Rapid Alleviation of African Animal Trypanosomosis by Integrated Disease Management Practices in the Agropastoral Zone of', *Tropical Animal Health and Production*, 31, pp. 89–102.

Baylis, M., Mbwabi, A. L. and Stevenson, P. (1994) 'The feeding success of tsetse flies , Glossina pallidipes (Diptera : Glossinidae), on oxen treated with pyrethroid pour-ons at Galana Ranch , Kenya', *Bulletin of Entomological Research*. University of Liverpool Library, 84, pp. 447–452.

Bekele, J. *et al.* (2010) 'Evaluation of Deltamethrin applications in the control of tsetse and trypanosomosis in the southern rift valley areas of Ethiopia', *Veterinary Parasitology*. Elsevier B.V., 168(3–4), pp. 177–184. doi: 10.1016/j.vetpar.2009.11.028.

Van Den Bossche, P. *et al.* (2004) 'A Large-scale Trial to Evaluate the Efficacy of a 1% Pour-on Formulation of Cyfluthrin (Cylence, Bayer) in Controlling Bovine Trypanosomosis in Eastern Zambia', *Tropical Animal Health and Production*, 36(1), pp. 33–43. doi: 10.1023/B:TROP.0000009519.18998.5b.

Van den Bossche, P. and de Deken, R. (2002) 'Seasonal variation in the distribution and abundance of the tsetse fly, Glossina morsitans morsitans in eastern Zambia', *Medical and Veterinary Entomology*, 16(2), pp. 170–176.

Brightwell, B. *et al.* (2001) 'Reality vs. rhetoric - A survey and evaluation of tsetse control in East Africa', *Agriculture and Human Values*, 18(2), pp. 219–233. doi: 10.1023/A:1011131826919.

Brightwell, R., Dransfield, R. D. and Williams, B. G. (1992) 'Factors affecting seasonal dispersal of the tsetse flies Glossina pallidipes and G. longipennis (Diptera: Glossinidae) at Nguruman, south-west Kenya', *Bulletin of Entomological Research*, 82(02), p. 167. doi: 10.1017/S0007485300051695.

Browning, D. M. *et al.* (2018) 'Integrating space and time: a case for phenological context in grazing studies and management', *Frontiers of Agricultural Science and Engineering*, 5(1), p. 44. doi: 10.15302/J-FASE-2017193.

Burgess, N. D. *et al.* (1992) 'Preliminary assessment of the distribution, status and biological importance of coastal forests in Tanzania', *Biological Conservation*, 62(3), pp. 205–218. doi: 10.1016/0006-3207(92)91048-W.

Bursell, E. (1961) 'Post-teneral development of the thoracic musculature in Tsetse flies', *Proceedings of the Royal Entomological Society of London (A)*, 36(4–6), pp. 69–74.

Bursell, E. (1966) 'The nutritional state of tsetse flies from different vegetation types in Rhodesia.', *Bulletin of entomological research*, 57(1), pp. 171–180. doi: 10.1017/S0007485300052810.

Bursell, E. (1970) 'Dispersal and concentration of Glossina', *The African Trypanosomiases*, pp. 382–394.

Carrasco-Escobar, G. *et al.* (2019) 'High-accuracy detection of malaria vector larval habitats using drone-based multispectral imagery', *PLOS Neglected Tropical Diseases*, 13(1), p. e0007105. doi: 10.1371/journal.pntd.0007105.

'CDC - African Trypanosomiasis - General Information - East African Trypanosomiasis FAQs' (no date). Available at: http://www.cdc.gov/parasites/sleepingsickness/gen_info/faqs-east.html (Accessed: 6 August 2015).

Chander, G. and Markham, B. (2003) 'Revised Landsat-5 TM Radiometrie Calibration Procedures and Postcalibration Dynamic Ranges', *IEEE Transactions on Geoscience and*

Remote Sensing, 41(11), p. 2674.

Chapuis, M. P. and Estoup, A. (2007) 'Microsatellite null alleles and estimation of population differentiation', *Molecular Biology and Evolution*, 24(3), pp. 621–631. doi: 10.1093/molbev/msl191.

Ciosi, M., Masiga, D. K. and Turner, C. M. R. (2014) 'Laboratory Colonisation and Genetic Bottlenecks in the Tsetse Fly Glossina pallidipes', *PLoS Neglected Tropical Diseases*, 8(2). doi: 10.1371/journal.pntd.0002697.

Cochard, R. and Edwards, P. J. (2011) 'Tree dieback and regeneration in secondary Acacia zanzibarica woodlands on an abandoned cattle ranch in coastal Tanzania', *Journal of Vegetation Science*, 22(3), pp. 490–502. doi: 10.1111/j.1654-1103.2011.01276.x.

Cochard, R., Edwards, P. J. and Weber, E. (2015) 'Post-ranching tree-grass interactions in secondary Acacia zanzibarica woodlands in coastal Tanzania - an experimental study', *Applied Vegetation Science*, 18(2), pp. 297–310. doi: 10.1111/avsc.12134.

Corander, J. *et al.* (2008) 'Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations.', *BMC bioinformatics*, 9, p. 539. doi: 10.1186/1471-2105-9-539.

Cornuet, J. M. and Luikart, G. (1996) 'Description and Power Analysis of Two Tests for Detecting Recent Population Bottlenecks From Allele Frequency Data', *Genetics*, 144, pp. 2001–2014. doi: Article.

Davies, H. (1964) 'The Eradication of Tsetse in the Chad River System of Northern Nigeria', *Journal of Applied Ecology*, 1(2), pp. 387–403.

Do, C. *et al.* (2014) 'NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data', *Molecular Ecology Resources*, 14(1), pp. 209–214. doi: 10.1111/1755-0998.12157.

Ducheyne, E. *et al.* (2009) 'The impact of habitat fragmentation on tsetse abundance on the plateau of eastern Zambia', *Preventive Veterinary Medicine*, 91(1), pp. 11–18. doi: 10.1016/j.prevetmed.2009.05.009.

Duguma, R. *et al.* (2015) 'Spatial distribution of Glossina sp. and Trypanosoma sp. in south-western Ethiopia', *Parasites and Vectors*. Parasites & Vectors, 8(1), pp. 1–10. doi: 10.1186/s13071-015-1041-9.

Dyer, N. a. *et al.* (2008) 'Molecular phylogenetics of tsetse flies (Diptera: Glossinidae) based on mitochondrial (COI, 16S, ND2) and nuclear ribosomal DNA sequences, with an emphasis on the palpalis group', *Molecular Phylogenetics and Evolution*. Elsevier Inc., 49(1), pp. 227–239. doi: 10.1016/j.ympev.2008.07.011.

Dyer, N. a. *et al.* (2011) 'Cryptic diversity within the major trypanosomiasis vector glossina fuscipes revealed by molecular markers', *PLoS Neglected Tropical Diseases*, 5(8). doi: 10.1371/journal.pntd.0001266.

Earl, D. A. and vonHoldt, B. M. (2012) 'STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method', *Conservation Genetics Resources*, 4(2), pp. 359–361. doi: 10.1007/s12686-011-9548-7.

Echodu, R. *et al.* (2013) ' Genetically Distinct Glossina fuscipes fuscipes Populations in the Lake Kyoga Region of Uganda and Its Relevance for Human African Trypanosomiasis ', *BioMed Research International*, 2013, pp. 1–12. doi: 10.1155/2013/614721.

Eisler, M. C. *et al.* (2003) 'Integrated control of vector-borne diseases of livestock - Pyrethroids: Panacea or poison?', *Trends in Parasitology*, 19(8), pp. 341–345. doi: 10.1016/S1471-4922(03)00164-8.

Esterhuizen, J. *et al.* (2005) 'Abundance and distribution of the tsetse flies, Glossina austeni and G. brevipalpis, in different habitats in South Africa', *Medical and Veterinary Entomology*, 19(4), pp. 367–371. doi: 10.1111/j.1365-2915.2005.00582.x.

Esterhuizen, J. *et al.* (2006) 'Selective use of odour-baited, insecticide-treated targets to control tsetse flies Glossina austeni and G. brevipalpis in South Africa', *Medical and Veterinary Entomology*, 20(4), pp. 464–469. doi: 10.1111/j.1365-2915.2006.00650.x.

Estes, A. B. *et al.* (2012) 'Land-cover change and human population trends in the greater Serengeti ecosystem from 1984–2003', *Biological Conservation*. Elsevier Ltd, 147(1), pp. 255–263. doi: 10.1016/j.biocon.2012.01.010.

Van Etten, J. (2011) 'Enzyme Polymorphism in Populations of the Tsetse Fly Glossina Pallidipes in Kenya', *Entomologia Experimentalis et Applicata*, 31(2–3), pp. 197–201. doi: 10.1111/j.1570-7458.1982.tb03135.x.

Felsenstein, J. (1989) 'PHYLIP - Phylogeny Inference Package (Version 3.2)', *Cladistics*, 5, pp. 164–166.

Fèvre, E. M. *et al.* (2005) 'A burgeoning epidemic of sleeping sickness in Uganda', *Lancet*,

366(9487), pp. 745–747. doi: 10.1016/S0140-6736(05)67179-6.

Fidelis Junior, O. L. *et al.* (2019) 'Comparison of conventional and molecular techniques for Trypanosoma vivax diagnosis in experimentally infected cattle', *Revista Brasileira de Parasitologia Veterinária*, 2961(june), pp. 203–209. doi: 10.1590/s1984-29612019034.

Ford, J., Nash, T. A. M. and Welch, J. R. (1970) 'Chapter 27: Control by Clearing of Vegetation', in *The African Trypanosomiases*, pp. 543–556.

Fornace, K. M. *et al.* (2018) 'Exposure and infection to Plasmodium knowlesi in case study communities in Northern Sabah, Malaysia and Palawan, The Philippines', *PLoS Neglected Tropical Diseases*, 12(6), pp. 1–16. doi: 10.1371/journal.pntd.0006432.

Fox, R. G. R. *et al.* (1993) 'Effect on herd health and productivity of controlling tsetse and trypanosomiasis by applying deltamethrin to cattle', *Tropical Animal Health and Production*, 25, pp. 203–214.

Franco, J. R. *et al.* (2014) 'Epidemiology of human African trypanosomiasis.', *Clinical Epidemiology*, 6, pp. 257–275.

Gallo, K. *et al.* (2011) 'Evaluation of the relationship between air and land surface temperature under clear- and cloudy-sky conditions', *Journal of Applied Meteorology and Climatology*, 50(3), pp. 767–775. doi: 10.1175/2010JAMC2460.1.

Gastellu Etchegorry, M. *et al.* (2001) 'Availability and affordability of treatment for human african trypanosomiasis', *Tropical Medicine and International Health*, 6(11), pp. 957–959. doi: 10.1046/j.1365-3156.2001.00764.x.

Geigy, R., Kauffmann, M. and Rogers, D. (1973) 'Sleeping Sickness Survey in the Serengeti Area (Tanzania) 1971', *Acta Tropica*, 30(1–2), pp. 13–48.

Giordani, F. *et al.* (2016) 'The animal trypanosomiases and their chemotherapy: A review', *Parasitology*, 143(14), pp. 1862–1889. doi: 10.1017/S0031182016001268.

GIS Geography (2019) *Planet Labs Imagery: The Entire Earth, Everyday - GIS Geography*. Available at: https://gisgeography.com/planet-labs-imagery/ (Accessed: 14 October 2019).

Glasgow, J. P. and Bursell, E. (1961) 'Seasonal variations in the fat content and size of Glossina swynnertoni Austen', *Bulletin of Entomological Research*, 51(4), pp. 705–713. doi: 10.1017/S0007485300057606.

Gondwe, N. *et al.* (2009) 'Distribution and density of tsetse flies (glossinidae: Diptera) at the

game/people/livestock interface of the nkhotakota game reserve human sleeping sickness focus in malawi', *EcoHealth*, 6(2), pp. 260–265. doi: 10.1007/s10393-009-0252-y.

Gooding, R. H. and Krafsur, E. S. (2005) 'Tsetse genetics: contributions to biology, systematics, and control of tsetse flies.', *Annual review of entomology*, 50(69), pp. 101–123. doi: 10.1146/annurev.ento.50.071803.130443.

Gooding, R. H., Solano, P. and Ravel, S. (2005) ' X-chromosome mapping experiments suggest occurrence of cryptic species in the tsetse fly Glossina palpalis palpalis ', *Canadian Journal of Zoology*, 82(12), pp. 1902–1909. doi: 10.1139/z05-002.

Goudet, J. (1995) 'FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics', *Journal of Heredity*, 86(6), pp. 485–486. doi: 10.1093/jhered/91.6.509.

Gouteux, J. P. *et al.* (1996) 'Traitement épicutané (Pour on et Spot on) du bétail contre Glossina fuscipes fuscipes en République centrafricaine', *Veterinary Research*, 27(3), pp. 273–284.

Hansen, M. C. *et al.* (2013) 'High-Resolution Global Maps of 21st-Century Forest Cover Change', *Science*, 342(November), pp. 850–854.

Hargrove, J. W. (1981) 'Discrepancies between estimates of tsetse fly populations using mark-recapture and removal trapping techniques', *Journal of applied ecology*, pp. 737–748. Available at: https://www.lib.byu.edu/cgi-bin/remoteauth.pl?url=http://search.ebscohost.com/login.aspx?direct=true&db=agr&AN=IND82056495&site=ehost-live&scope=site.

Hargrove, J. W. (1988) 'Tsetse: the limits to population growth.', *Medical and veterinary entomology*, 2(3), pp. 203–217. doi: 10.1111/j.1365-2915.1988.tb00184.x.

Hargrove, J. W. (1999) 'Nutritional levels of female tsetse Glossina pallidipes from artificial refuges', *Med Vet Entomol*, 13(2), pp. 150–164. doi: 10.1046/j.1365-2915.1999.00152.x.

Hargrove, J. W. (2000) 'A theoretical study of the invasion of cleared areas by tsetse flies (Diptera: Glossinidae).', *Bulletin of entomological research*, 90(3), pp. 201–209. doi: 10.1017/S0007485300000328.

Hargrove, J. W. *et al.* (2000) 'Insecticide-treated cattle for tsetse control : the power and the problems', *Med Vet Entomol*, 14, pp. 123–130.

Hargrove, J. W. (2003) 'Tsetse eradication : sufficiency , necessity and desirability'. DFID

Animal Health Programme.

Hargrove, J. W. *et al.* (2012) 'Modeling the control of trypanosomiasis using trypanocides or insecticide-treated livestock', *PLoS Neglected Tropical Diseases*, 6(5). doi: 10.1371/journal.pntd.0001615.

Hargrove, J. W. and Ackley, S. F. (2015) 'Mortality estimates from ovarian age distributions of the tsetse fly Glossina pallidipes Austen sampled in Zimbabwe suggest the need for new analytical approaches', *Bulletin of Entomological Research*, 105(03), pp. 294–304. doi: 10.1017/S0007485315000073.

Hargrove, J. W., Torr, S. J. and Kindness, H. M. (2003) 'Insecticide-treated cattle against tsetse (Diptera: Glossinidae): what governs success?', *Bulletin of Entomological Research*. University of Liverpool Library, 93(03), pp. 203–217. doi: 10.1079/BER2003234.

Hargrove, J. W. and Vale, G. A. (1980) 'Catches of Glossina morsitans morsitans Westwood and G. pallidipes Austen (Diptera: Glossinidae) in odour-baited traps in riverine and deciduous woodlands in the Zambesi Valley of Zimbabwe', *Bulletin of Entomological Research*, 70, pp. 571–578.

Hargrove, J. W. and Williams, B. G. (1995) 'A cost-benefit analysis of feeding in female tsetse', *Medical and Veterinary Entomology*, 9(2), pp. 109–119. doi: 10.1111/j.1365-2915.1995.tb00166.x.

Hartter, J. *et al.* (2016) 'Perceptions of risk in communities near parks in an African biodiversity hotspot', *Ambio*. Springer Netherlands, 45(6), pp. 692–705. doi: 10.1007/s13280-016-0775-8.

Hirt, C., Filmer, M. S. and Featherstone, W. E. (2010) 'Comparison and validation of the recent freely available ASTER-GDEM ver1, SRTM ver4.1 and GEODATA DEM-9s ver3 digital elevation models over Australia', *Australian Journal of Earth Sciences*, 57(3), pp. 337–347. doi: 10.1080/08120091003677553.

Holleley, C. E. and Geerts, P. G. (2009) 'Multiplex Manager 1.0: A cross-platform computer program that plans and optimizes multiplex PCR', *BioTechniques*, 46(7), pp. 511–517. doi: 10.2144/000113156.

Homewood, K. *et al.* (2002) 'Long-term changes in Serengeti-Mara wildebeest and land cover: Pastoralism, population, or policies?', *Proceedings of the National Academy of Sciences*, 98(22), pp. 12544–12549. doi: 10.1073/pnas.221053998.

Hyseni, C. *et al.* (2012) 'The population structure of Glossina fuscipes fuscipes in the Lake Victoria basin in Uganda: implications for vector control', *Parasites & Vectors*. Parasites & Vectors, 5(1), p. 222. doi: 10.1186/1756-3305-5-222.

Ismail, H. M. *et al.* (2016) 'Development of a Simple Dipstick Assay for Operational Monitoring of DDT', pp. 1–14. doi: 10.1371/journal.pntd.0004324.

Jackson, C. H. N. (1945) 'Comparative Studies of the Habitat Requirements of Tsetse Fly', *Journal of Animal Ecology*, 14(1), pp. 46–51.

Jackson, C. H. N. (1946) 'An artificially isolated generation of tsetse flies (Diptera)', *Bulletin of entomological research*, 37, pp. 291–299. doi: 10.1017/S0007485300022203.

Jackson, C. H. N. (1954) 'The hunger-cycles of Glossina morsitans Westwood and G. swynnertoni Austen', *Journal of Animal Ecology*, 23(2), pp. 368–372. doi: 10.1007/978-1-349-94186-5_232.

Kagbadouno, M. S. *et al.* (2011) 'Progress towards the eradication of Tsetse from the Loos islands, Guinea.', *Parasites & vectors*. BioMed Central Ltd, 4(1), p. 18. doi: 10.1186/1756-3305-4-18.

Kajunguri, D. *et al.* (2014) 'Modelling the Use of Insecticide-Treated Cattle to Control Tsetse and Trypanosoma brucei rhodesiense in a Multi-host Population', *Bulletin of Mathematical Biology*, 76(3), pp. 673–696. doi: 10.1007/s11538-014-9938-6.

Kalinowski, S. T., Wagner, A. P. and Taper, M. L. (2006) 'ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship', *Molecular Ecology Notes*, 6(2), pp. 576–579. doi: 10.1111/j.1471-8286.2006.01256.x.

Kamuanga, M. *et al.* (2001) 'Farmers' Perceptions of the Impact of Tsetse and Trypanosomosis Control on Livestock Production: Evidence from Southern Burkina Faso', *Tropical Animal Health and Production*, 33, pp. 141–153.

Kasilagila, G. (2003) 'Studies on trap effectiveness of tsetse flies (Glossina spp . (Diptera : Glossinidae)) in the Tanga Region of north eastern Tanzania', *Acta Tropica*, 87, pp. 385–392. doi: 10.1016/S0001-706X(03)00138-4.

Kgori, P. M., Modo, S. and Torr, S. J. (2006) 'The use of aerial spraying to eliminate tsetse from the Okavango Delta of Botswana', *Acta Tropica*, 99(2–3), pp. 184–199. doi: 10.1016/j.actatropica.2006.07.007.

Kimaro, E. G., Toribio, J. L. and Mor, S. M. (2017) 'Climate change and cattle vector-borne diseases: use of participatory epidemiology to investigate experiences in pastoral communities in northern Tanzania', *Preventive Veterinary Medicine*. Elsevier B.V. doi: 10.1016/j.prevetmed.2017.08.010.

Kitron, U. *et al.* (1996) 'Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS', *Journal of Animal Ecology*, 65(3), pp. 371–380. doi: 10.2307/5883.

Kofler, R., Schlötterer, C. and Lelley, T. (2007) 'SciRoKo: A new tool for whole genome microsatellite search and investigation', *Bioinformatics*, 23(13), pp. 1683–1685. doi: 10.1093/bioinformatics/btm157.

Koné, N. *et al.* (2011) 'Contrasting population structures of two vectors of African Trypanosomoses in Burkina Faso: Consequences for control', *PLoS Neglected Tropical Diseases*, 5(6), pp. 1–10. doi: 10.1371/journal.pntd.0001217.

Kopelman, N. *et al.* (2015) 'CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K', *Molecular Ecology Resources*, 15(5), pp. 1179–1191. doi: 10.1007/s10741-014-9462-7.Natural.

Krafsur, E. . *et al.* (1997) 'Breeding structure of Glossina pallidipes (Diptera:Glossinidea) populations in East and southern Africa', *Bulletin of Entomological Research*. YBP Library Services, 87, pp. 67–73.

Krafsur, E. S. *et al.* (2001) 'Genetic differentiation of Glossina morsitans centralis populations', *Insect Molecular Biology*, 10(4), pp. 387–395. doi: 10.2217/FON.09.6.Dendritic.

Krafsur, E. S. (2002a) 'Population structure of the tsetse fly Glossina pallidipes estimated by allozyme, microsatellite and mitochondrial gene diversities.', *Insect Molecular Biology*, 11(1), pp. 37–45. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

Krafsur, E. S. (2002b) 'Population structure of the tsetse fly Glossina pallidipes estimated by allozyme, microsatellite and mitochondrial gene diversities', *Insect Molecular Biology*, 11(1), pp. 37–45. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

Krafsur, E. S. and Endsley, M. a (2002) 'Microsatellite diversities and gene flow in the tsetse fly , Glossina morsitans s . l .', *Medical and Veterinary Entomology*, 16, pp. 292–300.

Krafsur, E. S., Marquez, J. G. and Ouma, J. O. (2016) 'Phylogeography and genealogy of the

tsetse fly Glossina pallidipes (Diptera: Glossinidae)', *International Journal of Tropical Insect Science*, 36(1), pp. 32–47. doi: 10.1017/S1742758415000223.

Krafsur, E. S. and Wohlford, D. L. (1999) 'Breeding structure of Glossina pallidipes populations evaluated by mitochondrial variation.', *The Journal of heredity*, 90(6), pp. 635–642. doi: 10.1093/jhered/90.6.635.

Kuepfer, I. *et al.* (2011) 'Clinical presentation of T.b. rhodesiense sleeping sickness in second stage patients from Tanzania and Uganda', *PLoS Neglected Tropical Diseases*, 5(3). doi: 10.1371/journal.pntd.0000968.

Langley, P. A. and Weidhaas, D. (1986) 'Trapping as a means of controlling tsetse , Glossina spp . (Diptera : Glossinidae): the relative merits of killing and of sterilization', *Bulletin of Entomological Research*, pp. 89–95.

Leak, S. G. A. (1995) 'A trial of a cypermethrin ' pour-on ' insecticide to control Glossina pallidipes , G . fuscipes fuscipes and G . tnorsitans submorsitans (Diptera : Glossinidae) in south-west Ethiopia', *Bulletin of Entomological Research*, 85, pp. 241–251.

Leak, S. G. A. (1998) *Tsetse Biology and Ecology*. CABI publishing.

Lehane, J. and Hargrove, J. W. (1988) 'Field experiments on a new method for determining age in tsetse flies (Diptera: Glossinidae)', *Ecological Entomology*, (13), pp. 319–322.

Lehane, M. J. *et al.* (2000) 'Grouping of trypanosome species in mixed infections in Glossina pallidipes', *Parasitology*, 120(6), pp. 583–592. doi: 10.1017/S0031182099005983.

Lehane, M. J. and Mail, T. S. (1985) 'Determining the age of adult male and female Glossina morsitans morsitans using a new technique', *Ecological Entomology*, 10(4), pp. 219–224.

Lejon, V., Jacobs, J. and Simarro, P. P. (2013) 'Elimination of sleeping sickness hindered by difficult diagnosis.', *Bulletin of the World Health Organization*, 91(10), p. 718. doi: 10.2471/BLT.13.126474.

Li, J. and Roy, D. P. (2017) 'A global analysis of Sentinel-2a, Sentinel-2b and Landsat-8 data revisit intervals and implications for terrestrial monitoring', *Remote Sensing*, 9(9). doi: 10.3390/rs9090902.

Lischer, H. E. L. and Excoffier, L. (2012) 'PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs', *Bioinformatics*, 28(2), pp. 298–299. doi: 10.1093/bioinformatics/btr642.

Lord, J. *et al.* (2018) 'Geostatistical models using remotely-sensed data predict savanna tsetse decline across the interface between protected and unprotected areas in Serengeti, Tanzania', *Journal of Applied Ecology*, pp. 1–11. doi: 10.5281/zenodo.1133429.

Lovis, L. *et al.* (2013) 'Determination of Acaricide Resistance in <I>Rhipicephalus (Boophilus) microplus</I> (Acari: Ixodidae) Field Populations of Argentina, South Africa, and Australia With the Larval Tarsal Test', *Journal of Medical Entomology*, 50(2), pp. 326–335. doi: 10.1603/me12127.

Luguru, S. M., Bennett, S. R. and Chizyuka, H. G. B. (1993) 'Observations on the Incidence of Bovine Trypanosomosis in Cattle Dipped in Deltamethrin in a Tsetse Infested Area of Zambia', *Tropical Animal Health and Production*, 25, pp. 129–130.

Lyimo, Z. C. *et al.* (2004) 'Post-partum reproductive performance of crossbred dairy cattle on smallholder farms in sub-humid coastal Tanzania', *Tropical Animal Health and Production*, 36(3), pp. 269–279. doi: 10.1023/B:TROP.0000016834.65290.d6.

Macdonald, G. (1956) 'Epidemiological basis of malaria control', *Bulletin of the World Health Organization*, pp. 613–626.

Magona, J. W., Walubengo, J. and Olaho-Mukani, W. (2004) 'Knowledge and attitudes of cattle owners regarding trypanosomosis control in tsetse-infested areas of Uganda', *Journal of the South African Veterinary Association*, 75(4), pp. 173–176. doi: 10.4102/jsava.v75i4.478.

Mahamat, M. H. *et al.* (2017) 'Adding tsetse control to medical activities contributes to decreasing transmission of sleeping sickness in the Mandoul focus (Chad)', (November 2013), pp. 1–16.

Malele, I. *et al.* (2003) 'The use of specific and generic primers to identify trypanosome infections of wild tsetse flies in Tanzania by PCR', *Infection, Genetics and Evolution*, 3(4), pp. 271–279. doi: 10.1016/S1567-1348(03)00090-X.

Malele, I. I. (2011) 'Fifty years of tsetse control in Tanzania: Challenges and prospects for the future', *Tanzania Journal of Health Research*, 13(5 SUPPL.ISS), pp. 1–10. doi: 10.4314/thrb.v13i5.9.

Malele, I., Nyingilili, H. and Msangi, A. (2011) 'Factors defining the distribution limit of tsetse infestation and the implication for livestock sector in Tanzania', *African Journal of Agricultural Research*, 6(10), pp. 2341–2347. doi: 10.5897/AJAR10.1050.

Manangwa, O. *et al.* (2019) ' Detecting Wahlund effects together with amplification problems: cryptic species, null alleles and short allele dominance in Glossina pallidipes populations from Tanzania ', *Molecular Ecology Resources*, pp. 0–1. doi: 10.1111/1755-0998.12989.

Matawa, F. *et al.* (2016) 'Modelling the spatial-temporal distribution of tsetse (Glossina pallidipes) as a function of topography and vegetation greenness in the Zambezi Valley of Zimbabwe', *Applied Geography*. Elsevier Ltd, 76, pp. 198–206. doi: 10.1016/j.apgeog.2016.09.008.

Mbassa, G. K. *et al.* (2017) *Analysing the Impact of Dip Construction and Rehabilitation and Acaricide Subsidy on Tick Borne Diseases in Tanzania*.

McCabe, C. T. and Bursell, E. (1975) 'Metabolism of digestive products in the tsetse fly, Glossina morsitans', *Insect Biochemistry*, 5, pp. 769–779.

Mélachio, T. T. T. *et al.* (2015) 'Effect of sampling methods, effective population size and migration rate estimation in Glossina palpalis palpalis from Cameroon', *Infection, Genetics and Evolution*, 33, pp. 150–157. doi: 10.1016/j.meegid.2015.04.023.

Meyer, A. *et al.* (2018) 'Integrated cost-benefit analysis of tsetse control and herd productivity to inform control programs for animal African trypanosomiasis', *Parasites and Vectors*. Parasites & Vectors, 11(1), pp. 1–14. doi: 10.1186/s13071-018-2679-x.

Meyrowitsch, D. W. *et al.* (2011) 'Is the current decline in malaria burden in sub-Saharan Africa due to a decrease in vector population?', *Malaria Journal*. BioMed Central Ltd, 10(1), p. 188. doi: 10.1186/1475-2875-10-188.

Minjauw, B. and Mcleod, A. (2003) 'Tick-borne diseases and poverty', *DFID Animal Health Programme, Centre for Tropical Veterinary Medicine*. Available at: https://www.cabdirect.org/cabdirect/abstract/20063155090.

Misana, S. B. (2012) 'Land-use/cover changes and their drivers on the slopes of Mount Kilimanjaro, Tanzania', *Journal of Geography and Regional Planning*, 5(6), pp. 151–164. doi: 10.5897/JGRP11.050.

Moore, A. and Waring, C. P. (2001) 'The effects of a synthetic pyrethroid pesticide on some aspects of reproduction in Atlantic salmon (Salmo salar L .).', 52, pp. 1–12.

Muhanguzi, D. *et al.* (2014) 'Improvements on Restricted Insecticide Application Protocol for Control of Human and Animal African Trypanosomiasis in Eastern Uganda', *PLoS*

Neglected Tropical Diseases, 8(10). doi: 10.1371/journal.pntd.0003284.

Muhanguzi, D. *et al.* (2015) 'Cost analysis of options for management of African Animal Trypanosomiasis using interventions targeted at cattle in Tororo District; South-eastern Uganda', *Parasites and Vectors*. Parasites & Vectors, 8(1), pp. 1–9. doi: 10.1186/s13071-015-0998-8.

Muller, G., Grébaut, P. and Gouteux, J. (2004) 'An agent-based model of sleeping sickness : simulation trials of a forest focus in southern Cameroon', 327, pp. 1–11. doi: 10.1016/j.crvi.2003.12.002.

Mwaseba, D. L. and Kigoda, K. J. (2017) 'Knowledge , attitude , and practices about tsetse control among communities neighbouring Serengeti', *Heliyon*. Elsevier Ltd., 3(e00324). doi: 10.1016/j.heliyon.2017.e00324.

Mweempwa, C. *et al.* (2015) 'Impact of habitat fragmentation on tsetse populations and trypanosomosis risk in Eastern Zambia.', *Parasites & vectors*. Parasites & Vectors, 8(1), p. 406. doi: 10.1186/s13071-015-1018-8.

Ngonyoka, A., Gwakisa, P. S., Estes, A. B., Salekwa, L. P., *et al.* (2017) 'Patterns of tsetse abundance and trypanosome infection rates among habitats of surveyed villages in Maasai steppe of northern Tanzania', *Infectious Diseases of Poverty*. Infectious Diseases of Poverty, 6(1), pp. 1–12. doi: 10.1186/s40249-017-0340-0.

Ngonyoka, A., Gwakisa, P. S., Estes, A. B., Nnko, H. J., *et al.* (2017) 'Variation of tsetse fly abundance in relation to habitat and host presence in the Maasai Steppe, Tanzania', *Journal of Vector Ecology*, 42(1), pp. 34–43. doi: 10.1111/jvec.12237.

Ngumbi, A. F. and Silayo, R. S. (2017) 'A cross-sectional study on the use and misuse of trypanocides in selected pastoral and agropastoral areas of eastern and northeastern Tanzania', *Parasites and Vectors*. Parasites & Vectors, 10(1), pp. 1–9. doi: 10.1186/s13071-017-2544-3.

Njoroge, M. M. *et al.* (2017) 'Exploring the potential of using cattle for malaria vector surveillance and control: a pilot study in western Kenya', *Parasites & Vectors*. Parasites & Vectors, 10(1), p. 18. doi: 10.1186/s13071-016-1957-8.

Nkya, R. *et al.* (2007) 'Constraints on smallholder market oriented dairy systems in the north eastern coastal region of Tanzania', *Tropical Animal Health and Production*, 39(8), pp. 627–636. doi: 10.1007/s11250-007-9059-4.

Nok, A. J. (2003) 'Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human African trypanosomiasis.', *Parasitology research*, 90(1), pp. 71–79. doi: 10.1007/s00436-002-0799-9.

Odiit, M. *et al.* (2005) 'Quantifying the level of under-detection of Trypanosoma brucei rhodesiense sleeping sickness cases', *Tropical Medicine and International Health*, 10(9), pp. 840–849. doi: 10.1111/j.1365-3156.2005.01470.x.

Odulaja, a *et al.* (2001) 'Spatial and temporal distribution of tsetse fly trap catches at Nguruman, southwest Kenya.', *Bulletin of entomological research*. University of Liverpool Library, 91(3), pp. 213–220. doi: 10.1079/BER200084.

Okello-Onen, J. *et al.* (1994) 'Control of Tsetse Flies in Uganda by Dipping Cattle in Deltamethrin', *Tropical Animal Health and Production*, 26, pp. 21–27.

Okeyo, W. A. *et al.* (2017) 'Temporal genetic differentiation in Glossina pallidipes tsetse fly populations in Kenya', *Parasites and Vectors*, 10(1), pp. 1–13. doi: 10.1186/s13071-017-2415-y.

Okiria, R. *et al.* (2002) 'Sustainability of tsetse control by subsequent treatment of 10% of a previously treated Ugandan cattle population with 1% w/v deltamethrin', *Tropical Animal Health and Production*, 34(2), pp. 105–114. doi: 10.1023/A:1014239305328.

Okiria, R. and Kalunda, M. (1994) 'Knock down and survival of tsetse flies fed on cattle and pigs dipped in deltamethrin', *Annals of Tropical Medicine and Parasitology*, 88(1), pp. 77–81. doi: 10.1080/00034983.1994.11812838.

Van Oosterhout, C. *et al.* (2004) 'MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data', *Molecular Ecology Notes*, 4(3), pp. 535–538. doi: 10.1111/j.1471-8286.2004.00684.x.

Orenge, C. O. *et al.* (2012) 'Trypanotolerance in N'Dama x Boran crosses under natural trypanosome challenge: Effect of test-year environment, gender, and breed composition', *BMC Genetics*, 13. doi: 10.1186/1471-2156-13-87.

Ouedraogo, I. *et al.* (2016) 'Land Cover Transition in Northern Tanzania', *Land Degradation and Development*, 27(3), pp. 682–692. doi: 10.1002/ldr.2461.

Ouma, J. O. *et al.* (2011) 'Genetic diversity and population structure of Glossina pallidipes in Uganda and western Kenya.', *Parasites & vectors*, 4, p. 122. doi: 10.1186/1756-3305-4-122.

Ouma, J. O., Marquez, J. G. and Krafsur, E. S. (2005) 'Macrogeographic population structure of the tsetse fly, Glossina pallidipes (Diptera: Glossinidae)', *Changes*, 95(5), pp. 437–447. doi: 10.1016/j.biotechadv.2011.08.021.Secreted.

Ouma, J. O., Marquez, J. G. and Krafsur, E. S. (2006) 'Microgeographical breeding structure of the tsetse fly, Glossina pallidipes in south-western Kenya', *Medical and Veterinary Entomology*, 20(1), pp. 138–149. doi: 10.1111/j.1365-2915.2006.00609.x.

Park, P. O. *et al.* (1972) 'A large-scale scheme for the eradication of Glossina morsitans morsitans Westw. in the Western Province of Zambia by aerial ultra-low-volume application of endosulfan', *Bulletin of Entomological Research*, 61(3), pp. 373–384. doi: 10.1017/S0007485300047271.

Pilson, R. D. and Pilson, B. M. (1967) 'Behaviour studies of Glossina morsitans Westw. in the field', *Bulletin of Entomological Research*, 57(2), pp. 227–257. doi: 10.1017/S0007485300049956.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000) 'Inference of Population Structure Using Multilocus Genotype Data', *Genetics*, 155, pp. 945–959.

Quinlan, R. J. and Gatehouse, A. G. (1981) 'Characteristics and implications of knockdown of the tsetse fly Glossina morsitans morsitans Westw. by deltamethrin', *Pesticide Science*, 12(4), pp. 439–442. Available at: file:///Y:/2820.pdf.

Randolph, S. E. and Rogers, D. J. (1978) 'Feeding cycles and flight activity in field populations of tsetse (Diptera : Glossinidae)', *Bulletin of Entomological Research*, 68, pp. 655–671.

Raphael, L. J. (2018) 'Vegetation Land Cover / Use Dynamics and Their Effects in Mbulu and Karatu Districts in the North- Eastern Highlands of Tanzania Vegetation Land Cover / Use Dynamics and Their Effects in Mbulu and', *International Journal of Geospatial and Environmental Research*, 5(1).

Raymond, M. and Rousset, F. (1995) 'GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism.', *Journal of Heredity*, 86, pp. 248–249.

Reed, D. N. *et al.* (2009) 'The spatial distribution of vegetation types in the Serengeti ecosystem: The influence of rainfall and topographic relief on vegetation patch characteristics', *Journal of Biogeography*, 36(4), pp. 770–782. doi: 10.1111/j.1365-2699.2008.02017.x.

Robinson, T., Rogers, D. and Williams, B. (1997a) 'Mapping tsetse habitat suitability in the common fly belt of southern Africa using multivariate analysis of climate and remotely sensed vegetation data.', *Medical and veterinary entomology*, 11(3), pp. 235–245.

Robinson, T., Rogers, D. and Williams, B. (1997b) 'Univariate analysis of tsetse habitat in the common fly belt of southern Africa using climate and remotely sensed vegetation data.', *Medical and veterinary entomology*, 11(3), pp. 223–34. doi: 10.1111/j.1365-2915.1997.tb00400.x.

Rock, K. S. *et al.* (2015) *Mathematical Models of Human African Trypanosomiasis Epidemiology*, *Advances in Parasitology*. Elsevier Ltd. doi: 10.1016/bs.apar.2014.12.003.

Rogers, D. (1979) 'Tsetse Population Dynamics and Distribution : A New Analytical Approach', *Journal of Animal Ecology*, 48(3), pp. 825–849.

Rogers, D. (1997) 'Study of a Natural Population of Glossina fuscipes fuscipes Newstead and a Model of Fly Movement Author (s): David Rogers Source : Journal of Animal Ecology , Vol . 46 , No . 1 (Feb ., 1977), pp . 309-330 Published by : British Ecological Society Stab', *Journal of Animal Ecology*, 46(1), pp. 309–330.

Rogers, D. J. (1988) 'A general model for the African trypanosomiases', *Parasitology*, 97(01), p. 193. doi: 10.1017/S0031182000066853.

Rogers, D. J. and Randolph, S. E. (1986) 'Distribution and Abundance of Tsetse Flies (Glossina Spp .)', *Journal of Animal Ecology*, 55(3), pp. 1007–1025.

Rogers, D. J. and Randolph, S. E. (1991) 'Mortality rates and population density of tsetse flies correlated with satellite imagery', *Nature*, 351, pp. 739–741.

Rowlands, G. J. *et al.* (1999) 'Estimating the effects of tsetse control on livestock productivity-a case study in southwest Ethiopia.', *Tropical animal health and production*, 31, pp. 279–294. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10509421.

Russell, T. L. *et al.* (2014) 'Evaluating the feasibility of using insecticide quantification kits (IQK) for estimating cyanopyrethroid levels for indoor residual spraying in Vanuatu.', *Malaria journal*, 13(1), p. 178. doi: 10.1186/1475-2875-13-178.

Schreiber, S. J. and Lloyd-Smith, J. O. (2009) 'Invasion dynamics in spatially heterogeneous environments.', *The American naturalist*, 174(4), pp. 490–505. doi: 10.1086/605405.

Sciarretta, a *et al.* (2010) 'Spatial clustering and associations of two savannah tsetse

species, Glossina morsitans submorsitans and Glossina pallidipes (Diptera: Glossinidae), for guiding interventions in an adaptive cattle health management framework.', *Bulletin of entomological research*, 100(6), pp. 661–70. doi: 10.1017/S0007485309990733.

Shaw, a. P. M. *et al.* (2013) 'Estimating the costs of tsetse control options: An example for Uganda', *Preventive Veterinary Medicine*. Elsevier B.V., 110(3–4), pp. 290–303. doi: 10.1016/j.prevetmed.2012.12.014.

Simarro, P. P. *et al.* (2010) 'The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases.', *International journal of health geographics*. BioMed Central Ltd, 9(1), p. 57. doi: 10.1186/1476-072X-9-57.

Simarro, P. P. *et al.* (2012) 'Estimating and Mapping the Population at Risk of Sleeping Sickness', *PLoS Neglected Tropical Diseases*, 6(10). doi: 10.1371/journal.pntd.0001859.

Sobrino, J. A., Jiminez-Munoz, J. C. and Paolini, L. (2004) 'Land surface temperature retrieval from LANDSAT TM 5', *Remote Sensing of Environment*, 90, pp. 434–440. doi: 10.1016/j.rse.2004.02.003.

Solano, P. *et al.* (1999) 'Intraspecific variability in natural populations of Glossina palpalis gambiensis from West Africa, revealed by genetic and morphometic analyses', *Medical and Veterinary Entomology*, 13, pp. 401–407.

Solano, P. *et al.* (2009) 'The population structure of Glossina palpalis gambiensis from Island and continental locations in coastal Guinea', *PLoS Neglected Tropical Diseases*, 3(3). doi: 10.1371/journal.pntd.0000392.

Squarre, D. *et al.* (2016) 'Human African Trypanosomiasis in the Kafue National Park, Zambia', *PLOS Neglected Tropical Diseases*, 10(5), p. e0004567. doi: 10.1371/journal.pntd.0004567.

Stanley, N. (2016) *Economic welfare analysis on tourism revenue distribution in Serengeti District, Tanzania*, *Sokoine University of Agriculture*.

Stanton, M. C. *et al.* (2018) 'The development of high resolution maps of tsetse abundance to guide interventions against human African trypanosomiasis in northern Uganda'. Parasites & Vectors, pp. 1–12.

Sungirai, M. *et al.* (2018) 'Genotyping acaricide resistance profiles of Rhipicephalus microplus tick populations from communal land areas of Zimbabwe', *Ticks and Tick-borne Diseases*. Elsevier, 9(1), pp. 2–9. doi: 10.1016/j.ttbdis.2017.10.017.

Swallow, B. M. (2000) 'Impacts of trypanosomiasis on African agriculture.', *PAAT technical and scientific series (FAO)*, 2, pp. 1–46. doi: 10.1371/journal.pone.0081950.

Swallow, B. M., Mulatu, W. and Leak, S. G. A. (1995) 'Potential demand for a mixed public-private animal health input : evaluation of a pour-on insecticide for controlling tsetse-transmitted trypanosomiasis in Ethiopia', *Preventive Veterinary Medicine*, 24, pp. 265–275.

Symeonakis, E., Robinson, T. and Drake, N. (2007) 'GIS and multiple-criteria evaluation for the optimisation of tsetse fly eradication programmes', *Environmental Monitoring and Assessment*, 124(1–3), pp. 89–103. doi: 10.1007/s10661-006-9210-0.

Tamura, K. *et al.* (2013) 'MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0', *Molecular biology and evolution*, 30(12), pp. 2725–2729.

Tanga, Tanzania Weather Averages | Monthly Average High and Low Temperature | World Weather Online (no date). Available at: https://www.worldweatheronline.com/tanga-weather-averages/tanga/tz.aspx (Accessed: 25 June 2019).

Tanner, M. *et al.* (2015) 'Bovine Tuberculosis and Brucellosis in Cattle and African Buffalo in the Limpopo National Park, Mozambique', *Transboundary and Emerging Diseases*, 62(6), pp. 632–638. doi: 10.1111/tbed.12210.

Thompson, J. W. *et al.* (1991) 'Studies on the efficacy of deltamethrin applied to cattle for the control of tsetse flies (Glossina spp.) in southern Africa', *Tropical Animal Health and Production*, 23(4), pp. 221–226. doi: 10.1007/BF02357104.

Thomson, M. C. (1987) 'The effect on tsetse flies (glossina spp.) of deltamethin applied to cattle either as a spray or incorporated into ear-tags', *Tropical Pest Management*, 33(4), pp. 329–335. doi: 10.1080/09670878709371177.

Thumbi, S. M. *et al.* (2014) 'Parasite co-infections and their impact on survival of indigenous cattle', *PLoS ONE*, 9(2). doi: 10.1371/journal.pone.0076324.

Tobler, M. W., Cochard, R. and Edwards, P. J. (2003) 'The impact of cattle ranching on large-scale vegetation patterns in a coastal savanna in Tanzania', *Journal of Applied Ecology*, 40(3), pp. 430–444. doi: 10.1046/j.1365-2664.2003.00816.x.

Tong, J. *et al.* (2011) 'Challenges of controlling sleeping sickness in areas of violent conflict: Experience in the Democratic Republic of Congo', *Conflict and Health*, 5(1), pp. 1–8. doi: 10.1186/1752-1505-5-7.

Torr, S. J. (1985) 'The susceptibility of Glossina pallidipes Austen (Diptera: Glossinidae) to insecticide deposits on targets', *Bulletin of Entomological Research*, 75(3), pp. 451–458. doi: 10.1017/S0007485300014553.

Torr, S. J. *et al.* (1997) 'Methods for dispensing odour attractants for tsetse flies (Diptera: Glossinidae)', *Bulletin of Entomological Research*, 87(03), p. 299. doi: 10.1017/S0007485300037251.

Torr, S. J. *et al.* (2007) 'Is there safety in numbers? The effect of cattle herding on biting risk from tsetse flies', *Medical and Veterinary Entomology*, 21(4), pp. 301–311. doi: 10.1111/j.1365-2915.2007.00705.x.

Torr, S. J. *et al.* (2011) 'Responses of tsetse flies, Glossina morsitans morsitans and Glossina pallidipes, to baits of various size', *Medical and Veterinary Entomology*, 25(4), pp. 365–369. doi: 10.1111/j.1365-2915.2011.00947.x.

Torr, S. J. and Hargrove, J. W. (1998) 'Factors affecting the landing and feeding responses of the tsetse fly Glossina pallidipes to a stationary ox', *Medical and Veterinary Entomology*, 12, pp. 196–207.

Torr, S. J., Hargrove, J. W. and Vale, G. a (2005) 'Towards a rational policy for dealing with tsetse.', *Trends in parasitology*, 21(11), pp. 537–41. doi: 10.1016/j.pt.2005.08.021.

Torr, S. J., Kindness, H. M. and Obsomer, V. (2000) *Cattle management practices in tsetse-affected areas* .

Torr, S. J., Mangwiro, T. N. C. and Hall, D. R. (2011) 'Shoo fly, don't bother me! Efficacy of traditional methods of protecting cattle from tsetse', *Medical and Veterinary Entomology*, 25(2), pp. 192–201. doi: 10.1111/j.1365-2915.2010.00942.x.

Torr, S. J., Maudlin, I. and Vale, G. a. (2007) 'Less is more: Restricted application of insecticide to cattle to improve the cost and efficacy of tsetse control', *Medical and Veterinary Entomology*, 21(1), pp. 53–64. doi: 10.1111/j.1365-2915.2006.00657.x.

Torr, S. J., Maudlin, I. and Vale, G. A. (2007) 'Less is more: Restricted application of insecticide to cattle to improve the cost and efficacy of tsetse control', *Medical and Veterinary Entomology*, 21(1), pp. 53–64. doi: 10.1111/j.1365-2915.2006.00657.x.

Torr, S. J., Parker, A. G. and Leigh-Browne, G. (1989) ' The responses of Glossina pallidipes Austen (Diptera: Glossinidae) to odourbaited traps and targets in Somalia ', *Bulletin of Entomological Research*, 79(1), pp. 99–108. doi: 10.1017/s0007485300018605.

Torr, S. J. and Vale, G. A. (2011) 'Is the even distribution of insecticide-treated cattle essential for tsetse control? modelling the impact of baits in heterogeneous environments', *PLoS Neglected Tropical Diseases*, 5(10). doi: 10.1371/journal.pntd.0001360.

Treydte, A. C., Edwards, P. J. and Suter, W. (2005) 'Shifts in native ungulate communities on a former cattle ranch in Tanzania', *African Journal of Ecology*, 43(4), pp. 302–311. doi: 10.1111/j.1365-2028.2005.00573.x.

USGS EarthExplorer - Home (no date). Available at: https://earthexplorer.usgs.gov/ (Accessed: 1 July 2019).

Vale, G. *et al.* (1988) 'Odour-baited targets to control tsetse flies, Glossina spp. (Diptera: Glossinidae) in Zimbabwe', *Bulletin of Entomological Research*, 78, pp. 31–49.

Vale, G. a. (1974) 'The responses of tsetse flies (Diptera, Glossinidae) to mobile and stationary baits', *Bulletin of Entomological Research*, 64(04), p. 545. doi: 10.1017/S0007485300035860.

Vale, G. A., Hursey, B. S., Hargove, J. W., *et al.* (1984) 'The use of small plots to study populations of tsetse (Diptera: Glossinidae): Difficulties associated with population dispersal', *International Journal of Tropical Insect Science*, 5(05), pp. 403–410.

Vale, G. A., Hursey, B. S., Hargrove, J. W., *et al.* (1984) 'The use of small plots to study populations of tsetse (Diptera: Glossinidae)', 5(5), pp. 403–410.

Vale, G. A. *et al.* (2015) 'Optimal Strategies for Controlling Riverine Tsetse Flies Using Targets : A Modelling Study', *PLoS neglected tropical diseases*, 9(3), pp. 1–21. doi: 10.1371/journal.pntd.0003615.

Vale, Glyn A. *et al.* (2015) 'Pyrethroid Treatment of Cattle for Tsetse Control: Reducing Its Impact on Dung Fauna', *PLoS Neglected Tropical Diseases*, 9(3), pp. 1–11. doi: 10.1371/journal.pntd.0003560.

Vale, G. A. and Hall, D. R. (1985) 'The role of 1-octen-3-ol, acetone and carbon dioxide in the attraction of tsetse flies, Glossina spp. (Diptera: Glossinidae), to ox odour', *Bulletin of Entomological Research*, 75(2), pp. 209–218. doi: 10.1017/S0007485300014292.

Vale, G. A., Mutika, G. and Lovemore, D. F. (1999) 'Insecticide-treated cattle for controlling tsetse flies (Diptera: Glossinidae): some questions answered, many posed', *Bulletin of Entomological Research*, 89(6), pp. 569–578. doi: 10.1017/S0007485399000723.

Vale, G. A. and Torr, S. J. (2005) 'User-friendly models of the costs and efficacy of tsetse control: Application to sterilizing and insecticidal techniques', *Medical and Veterinary Entomology*, 19(3), pp. 293–305. doi: 10.1111/j.1365-2915.2005.00573.x.

Vickerman, K. (1985) 'Developmental Cycles and Biology of Pathogenic Trypanosomes', *British Medical Bulletin*, 41(2), pp. 105–114. Available at: http://bmb.oxfordjournals.org/content/41/2/105.full.pdf.

Vieira, M. L. C. *et al.* (2016) 'Microsatellite markers: What they mean and why they are so useful', *Genetics and Molecular Biology*, 39(3), pp. 312–328. doi: 10.1590/1678-4685-GMB-2016-0027.

Vreysen, M. J. B. *et al.* (2000) 'Glossina austeni (Diptera: Glossinidae) Eradicated on the Island of Unguja, Zanzibar, Using the Sterile Insect Technique', *Journal of Economic Entomology*, 93(1), pp. 123–135. doi: 10.1603/0022-0493-93.1.123.

Vreysen, M. J. B. *et al.* (2013) 'Release-Recapture Studies Confirm Dispersal of Glossina palpalis gambiensis between River Basins in Mali', *PLoS Neglected Tropical Diseases*, 7(4), pp. 5–8. doi: 10.1371/journal.pntd.0002022.

Vudriko, P. *et al.* (2016) 'Emergence of multi-acaricide resistant Rhipicephalus ticks and its implication on chemical tick control in Uganda', *Parasites and Vectors*. Parasites & Vectors, 9(1). doi: 10.1186/s13071-015-1278-3.

Wardrop, N. a *et al.* (2013) 'The dispersal ecology of Rhodesian sleeping sickness following its introduction to a new area.', *PLoS neglected tropical diseases*, 7(10), p. e2485. doi: 10.1371/journal.pntd.0002485.

Warnes, M. L. *et al.* (1999) 'Evaluation of insecticide-treated cattle as a barrier to re-invasion of tsetse to cleared areas in northeastern Zimbabwe', *Medical and Veterinary Entomology*, 13(2), pp. 177–184. doi: 10.1046/j.1365-2915.1999.00148.x.

Welburn, S. C. *et al.* (2006) 'Crisis, what crisis? Control of Rhodesian sleeping sickness', *Trends in Parasitology*, 22(3), pp. 123–128. doi: 10.1016/j.pt.2006.01.011.

Welburn, S. C., Molyneux, D. H. and Maudlin, I. (2015) 'Beyond Tsetse – Implications for Research and Control of Human African Trypanosomiasis Epidemics', *Trends in Parasitology*. Elsevier Ltd, xx(ii), pp. 1–12. doi: 10.1016/j.pt.2015.11.008.

Wellde, B. T. *et al.* (1989) 'Review of tsetse control measures taken in the Lambwe Valley in 1980-1984.', *Annals of Tropical Medicine and Parasitology*, 83(1), pp. 119–25. doi:

10.1080/00034983.1989.11812416.

WHO (2018) *Distribution of human African trypanosomiasis, worldwide, 2017*. Available at: http://gamapserver.who.int/mapLibrary/Files/Maps/HAT_rh_2017.png (Accessed: 3 July 2019).

WHO (no date a) *GHO | By category | Number of new reported cases (T.b. gambiense) - Data by country*, *WHO*. World Health Organization. Available at: http://apps.who.int/gho/data/node.main.A1636?lang=en (Accessed: 16 June 2019).

WHO (no date b) *GHO | By category | Number of new reported cases (T.b. rhodesiense) - Data by country*, *WHO*. World Health Organization. Available at: http://apps.who.int/gho/data/node.main.A1637?lang=en (Accessed: 16 June 2019).

Williams, B., Brightwell, R. and Dransfield, R. (1990) 'Monitoring tsetse fly populations.:II. The effect of climate on trap catches of Glossina pallidipes', *Medical and Veterinary Entomology*, 4(2), pp. 181–193. doi: 10.1111/j.1365-2915.1990.tb00276.x.

Williamson, D. L. *et al.* (1983) 'Integration of insect sterility and insecticides for control of Glossina morsitans morsitans Westwood (Diptera : Glossinidae) in Tanzania . II . Methods of sterilisation , transportation and release of sterilised males', *Bulletin of Entomological Research*. University of Liverpool Library, 73, pp. 267–273.

World Bank Group (2015) *Tanzania's Tourism Futures: Harnessing Natural Assets*. Available at: http://documents.worldbank.org/curated/en/204341467992501917/pdf/96150-REVISED-PN-P150523-PUBLIC-Box393206B.pdf (Accessed: 18 June 2019).

Ye, J. *et al.* (2012) 'Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction', *BMC bioinformatics*, 13, p. 134.

Yunta, C. *et al.* (2016) 'Pyriproxyfen is metabolized by P450s associated with pyrethroid resistance in An. gambiae', *Insect Biochemistry and Molecular Biology*, 78, pp. 50–57. doi: 10.1016/j.ibmb.2016.09.001.

Appendix

A.1 HPLC Standards

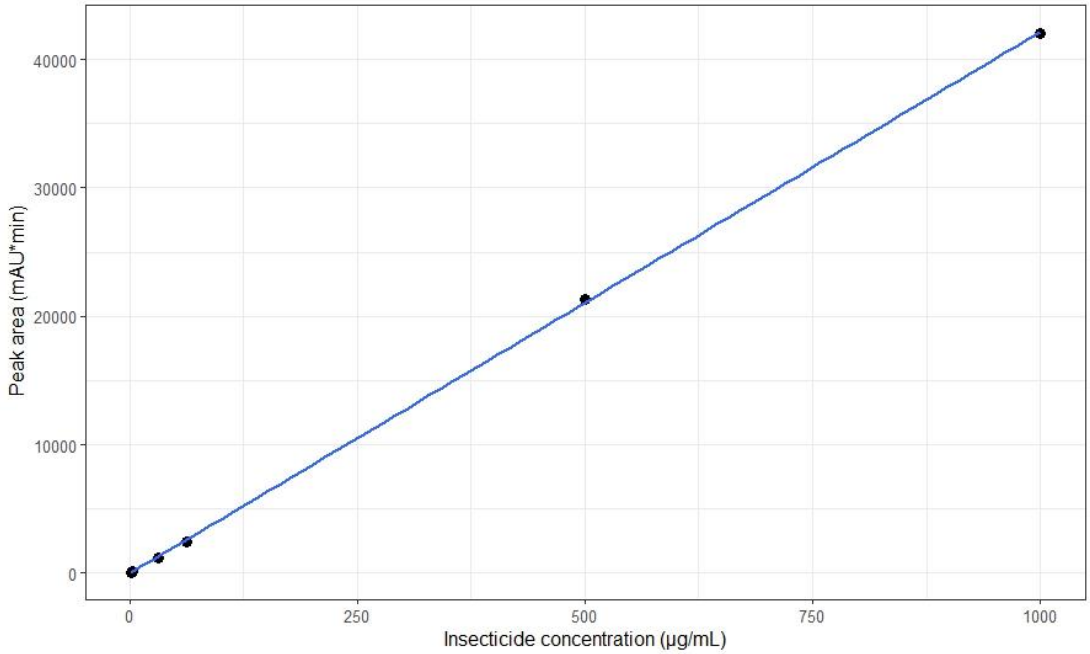

Figure 75. The area under the peaks for HPLC standard samples containing alpha-cypermethrin.

A.2 Log-CPO score and model coefficients from Lord *et al*., (2018) model

Table 47. Negative of the sum of the log conditional predictive ordinates (log-CPO score) for various linear models tested with catch data and environmental variable values from the Serengeti National Park (Adapted from Lord et al., (2018)).

Environmental variable	Log-CPO score
Band 7	48.9
NDVI	59.6
LST	57.8
Elevation	59.8
Band 7, elevation	47.8
Band 7, elevation, NDVI	47.1
Band 7, elevation, LST	47.8
Band 7, elevation, NDVI, LST	32.7

Table 48. Posterior estimates of model coefficients from the model with the highest log-CPO score.

Serengeti Model coefficients	Mean
Band 7	-43.10
NDVI	-7.633
Elevation	-0.003
LST	0.074
Intercept	12.42
Range	3,323

A.3 R and QGIS code for satellite image processing and model analysis

A.3.1 Landsat 4-5 conversion to radiance

```
rad.calc <- function(file){
 lmax <- 14.38
 lmin <- -0.15
 qcalmax <- 255
  input.raster <- raster(file)

 spec.rad <- ((lmax - lmin) / qcalmax)*input.raster + lmin
  return(spec.rad)
}
```

A.3.2 Landsat 4-5 conversion to reflectance

```
#pi * radiance (input raster) * earth-sun distance in astronomical units^2/ mean solar exoatmospheric irradiances then get sun angle from meta file

reflect.calc.45 <- function(file,sunang){
  input.raster <-file
  ref <- (pi*input.raster*(0.9840*0.9840))/(80.67*sin(pi*sunang/180))
return(ref)
}
```

A.3.3 Landsat 8 conversion to reflectance

```
reflect.calc.8 <- function(file,meta.d){    # arguments are the input raster file and the data extracted from the metadata
    ref.add <- meta.d$ref.add
  ref.mult <- meta.d$ref.mult
  sun.ang <- meta.d$sunelev
  input.raster <- raster(file)
    toa <- ref.mult*input.raster+ref.add     # I think refer to Yale paper and USGS website for top of atmosphere calculation
```

```
  toa.sunangle.corr <- toa/(sin(sun.ang))
 return(toa.sunangle.corr)
 }
```

A.3.4 Using coefficients from Lord *et al.*, (2018) model to create a predictive map for Tanga

```
coefb7 <- -43.10
coeftemp <- 0.074
coefndvi <- -7.633
coefelev <- -0.003
intercept <- 12.42

b7 <- raster("b7_coast_500m.tif")
temp <- raster("temp_coast_500m.tif")
elev <- raster("elev_coast_500m.tif")
ndvi <- raster("ndvi_coast_500m.tif")

predict <- intercept + coefb7*as.matrix(b7) + coeftemp*as.matrix(temp) +
coefndvi*as.matrix(ndvi)
+ coefelev*as.matrix(elev)

predict <- raster(predict)
crs(predict) <- crs(b7)
extent(predict) <- extent(b7)
writeRaster(10^predict,filename="serengeti_predict_coast ",format="GTiff")
```

A.3.5 The structure of the Lord *et al.*, (2018) generalised linear geostatistical model used with the new data collected in Tanga, with an example list of covariates used.

```
elev <- raster("elev_coast_500m.tif")
b7 <- raster("b7_coast_500m.tif")
ndvi <- raster("ndvi_coast_500m.tif")
temp <- raster("temp_coast_500m.tif")
covlist <- list(b7=b7,ndvi=ndvi
            ,temp=temp, elev=elev)

gp.saaboth.fit <- glgm(formula=log10(av_count+1) ~
ndvi+b7+temp+elev,data=gp.counts,grid=ndvi,covariates=covlist,family="gaussian",
                priorCI=list(sd=c(0.01,5),range=c(500,5000),sdNugget=c(0.01,5)),
                control.compute=list(dic=T,cpo=T))
```

A.3.6 Raster calculator in QGIS

```
Step 1: ("ndvi@1" <= 0.5 AND  “ndvi@1” >=0.2) *( 0.004 * ( "ndvi@1" - 0.2 / 0.5 - 0.2)^2 +
0.986)
Step 2:("ndvi@1" > 0.5) * 0.99 + ("ndvi@1" < 0.2) * 0.973 + ("ndvi@1" >= 0.2 AND
"ndvi@1"<= 0.5 ) * "emissivity_step1@1"
Step 3: temperature raster / ( 1 + ( 10.8 * b / 14380 ) * log(emissivity raster)).
```

A. 4 Mini questionnaire for individual animal sampling

Date: .....

Village:..... Subvillage:.....

Household Code:.....

Is insecticide used on the whole herd?		
Y	N	If N, describe how many have insecticide on...

	ID	Date last treated with insecticide	Name of insecticide used & method of application	Where was insecticide applied on the animal?	Location of hair sample/s
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					

12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					

A.5 Dip tank record from February 2017

Dip tanks Serengeti District 9/2/17

TAARIFA YA HALI YA MAJOSHO WILAYANI

A. MAJOSHO MAZIMA YANAYOOGESHA WAKATI WOTE

NO.	JINA LA JOSHO	SABABU YA KUTOOGESHA
1	NYAMISINGISI	
2	SINGISI	
3	NATTA	
5	TABORA B	
6	ROBANDA	
7	PARKNYIGOTI	
8	MAJIMOTO	
9	BWITENGI	
10	KONO	
11	KWITETE	
12	GUSUHI	
13	MERENGA	
14	NYAMOKO	
15	NYAMATOKE	
B. MAJOSHO YANAYOOGESHA KWA KUSIMAMASIMAMA		
16	NYAKITONO	CHANZO CHA MAJI NI MTO NA MAJI YAKIKAUKA UOGESHAJI HUSIMAMA
17	NYIBOKO	CHANZO CHA MAJI NI MTO NA MAJI YAKIKAUKA UOGESHAJI HUSIMAMA
18	BUCHANCHARI	CHANZO CHA MAJI NI MTO NA MAJI YAKIKAUKA UOGESHAJI HUSIMAMA
19	MARASOMOCHE	CHANZO CHA MAJI NI MTO NA MAJI YAKIKAUKA UOGESHAJI HUSIMAMA
20	NYICHOKA	CHANZO CHA MAJI NI BWAWA DOGO LINAKAUKA KIANGAZI
C. MAJOSHO AMBAYO YANAMGOGORO WA ARDHI		
21	MONUNA	WALIOTOA ENEO VIJIJI VIMESHINDWA KUWAFIDIA ENEO NA WALIOTOA MAENEO YAO WANALIMA HADI JOSHONI
22	GEASRYA	
C. MAJOSHO MABOVU YANAYOHITAJI UKARABATI MDOGO		
23	HEKWE	KUNA NYUFA KUTA ZA SHIMO
24	MOSONGO	KUNA NYUFA KUTA ZA SHIMO
25	KITUNGURUMA	KUNA NYUFA KUTA ZA SHIMO
26	MUSATI	KUNA NYUFA KUTA ZA SHIMO
27	KENYAMONTA	KUNA NYUFA KUTA ZA SHIMO
28	KYAMBAHI	KUNA NYUFA KUTA ZA SHIMO
29	NYANSURURA	KUNA NYUFA KUTA ZA SHIMO
30	NYAMITITA	KUNA NYUFA KUTA ZA SHIMO
D. MAJOSHO YANAYOHITAJI UKARABATI MKUBWA AU KUJENGWA UPYA		
31	MAGANGE	UKUTA UMEBOMOKA NA SHIMO LINA NYUFA
32	NYAGASENSE	UKUTA UMEBOMOKA NA SHIMO LINA NYUFA
33	KISANGURA	UKUTA UMEBOMOKA NA SHIMO LINA NYUFA
34	RUNG'ABURE	UKUTA UMEBOMOKA NA SHIMO LINA NYUFA
35	KENYANA	KUTA NA PAA VIPO LAKINI SHIMO LINAVUJA
36	KEBOSONGO	KUTA NA PAA VIPO LAKINI SHIMO LINAVUJA
E. MAJOSHO AMBAYO MAENEO YALIPO KWA SASA NI MAKAZI AU MASHAMBA		
37	MUGUMU	LIMEZUNGUKWA NA MAKAZI
38	KEBANCHABANCHA	LIMEZUNGUKWA NA MASHAMBA

And working

water restrictions

land issues

need minor repairs

in complete disrepair

areas now residential

A.6 Ethical clearance for VVBDRI sampling

Ms Rachel Hopper
Liverpool School of Tropical Medicine
Pembroke Place
Liverpool
L3 5QA

LSTM
LIVERPOOL SCHOOL OF TROPICAL MEDICINE
Pembroke Place,
Liverpool, L3 5QA, UK
Tel: +44(0)151 705 3100
Fax: +44(0)151 705 3370
www.lstmed.ac.uk

Tuesday, 17 April 2018

Dear Ms Hopper,

Re. Research Protocol (18-032) Cattle sampling to detect presence of trypanosomes in a herd belonging to the Vector and Vector-Borne Disease Research Institute, Tanzania.

I am pleased to confirm that LSTM supports the above-mentioned research study. It is also acknowledged that LSTM REC approval is exempted, given that cattle alone are the focus of this research project.

Please note that LSTM support is conditional upon compliance with the relevant regulatory requirements.

All study staff should be given the appropriate training, relevant to their responsibilities as defined within the study protocol.

LSTM Research Governance and Ethics Office should receive annual study progress and final close out reports via lstmgov@lstmed.ac.uk

Yours Sincerely,

Carl Henry
Research Governance Manager
Research Governance and Ethics Office

A.7 SRUC ethical clearance for Serengeti cattle sampling

Borderline experiments for review (before any requirement to fill in AEC forms)

It is a requirement that all SRUC work involving animals is notified to SRUC's AEC. This form is aimed to provide information to enable an AEC view on work that SRUC is only marginally involved with (e.g. at third party premises, or is only mildly interventionist (e.g. a largely observation study, or using data already being collected for other purposes), or primarily educational (an undergraduate/MSc study), or it is uncertain whether animal ethics is involved (work in slaughterhouse). There are three potential outcomes:

a) *If the work is not judged to be research/experimental then submitter will be thanked for notification and staff/student provided with AEC view that work can go ahead without further involvement of AEC*
b) *information on the form is sufficient for an approval process by AEC to go ahead*
c) *the work described requires a full experimental appraisal form.*

Please give brief details, sufficient to enable each of these cases to be judged.

---

1) Supervisor's or Principle Member of Staff's Details:

Name(s): Harriet Auty

Institution(s): SRUC Research

Tel(s): 01463 246071

Email(s): harriet.auty@sruc.ac.uk

If applicable, student's name: NA

2) Title of proposed study: Cross sectional survey of *Trypanosoma spp.* prevalence in livestock in Tanzania (part of a BBSRC-funded project led by

LSTM, entitled “Tackling human African trypanosomiasis on the edge of protected areas”)

3) **Proposed location of study:** Serengeti District, Tanzania

4) **Proposed start and end dates of study:** July 2016-August 2016

If this is a pilot trial, that will not distress the animals involved, then the above detail PLUS ITEMS 7-8 BELOW is sufficient for submission to the AEC for assessment. Examples would be testing new pedometers, or assessing if a radio collar works. Please tick this box if you think this applies to this trial. Do not fill in any more of this form until requested to do so. ☐

---

5) **Hypothesis:** Cattle in this area are (i) carrying *Trypanosoma brucei* s.l., which causes human African trypanosomiasis, and (ii) carrying trypanosome species that cause livestock disease.

6) **Brief background:** Both cattle and wildlife can carry *Trypanosoma brucei rhodesiense*, which causes the human disease human African trypanosomiasis (HAT), also known as sleeping sickness. The pathogen is carried by a vector, the tsetse fly. Tsetse are widespread within protected areas such as national parks where the wildlife act as reservoir hosts and maintain transmission of the disease. A major outstanding question about HAT transmission is what role cattle play around protected areas. It is know that they can carry the human pathogen, alongside other trypanosome species which are pathogenic to cattle. This study aims to quantify the prevalence of *T. brucei* s.l. in cattle, and obtain isolates for further genetic analysis. The prevalence of other trypanosome species, which cause disease in cattle, and of tick-borne disease in cattle (East Coast Fever) will also be determined. Trypanosomiasis and East Coast Fever are the most significant diseases of cattle in this area.

7) Summary of study design:

Herd selection: The study is an observational randomised cross sectional survey of disease prevalence in cattle. To select herds for sampling, villages will be selected from those where the village centre is located within 5km of the protected area boundary, within Serengeti District. Within a village, subvillages will be randomly selected, and a number of households that keep cattle will be randomly selected from within these subvillages. At each level (village and subvillage) the aims of the study will be explained to local leaders by the TTRI veterinarian (see section 8) and a local District Livestock Officer, in the local language (Swahili). Once a household is selected, the TTRI veterinarian and the local District Livestock Officer, will explain the aim of the study to the head of the household, and obtain written permission to sample their cattle, and to ask them some questions about the animals.

Sampling: All sampling will be coordinated by the TTRI veterinarian, accompanied by an experienced veterinary technician and a local District Livestock Officer. Each herd is usually contained within a thorn-fenced boma (enclosure). 20-30 cattle will be randomly selected within each herd (all animals if the herd size is less than this). For each animal, the animal will be restrained with ropes as no facilities for animal handling are available. This is the usual method for cattle restraint in this area. Local people who have experience of handling cattle in this way will assist with animal handling. A blood sample will be collected by jugular veno-puncture using a vacutainer and needle. One aliquot will be preserved in a Paxgene tube and a second processed to obtain serum. In addition, in a subset of animals, a peripheral capillary sample will be collected from an ear vein, using a sterile lancet and capillary tube (in the same way that finger prick samples are taken from people). This technique increases the sensitivity of detecting *T. brucei*, due to its tendency to localise in peripheral tissues, and is widely used in trypanosomiasis surveys. A hair sample from the animal will also be collected, to quantify insecticide presence. The animal will be marked to ensure it is not sampled again, then released back into the herd. The herd owner will be asked to confirm the animal's sex, age, and any treatments, prophylactic drugs, vaccines or insecticide it has received in the previous 6 months. The TTRI veterinarian will provide veterinary advice on any animals where this is necessary, including treatment for trypanosomiasis if required.

Analysis: The blood samples collected will be exported to the Roslin Institute, where they will undergo molecular analysis to diagnose the presence of human and livestock trypanosome species, as well as tick borne-pathogens.

Results and feedback: Feedback from the analysis will be provided to the original herd owners, and to local livestock stakeholders, at a future date, as well as information about how they can optimise disease control.

8) **Where SRUC is not the main or sole initiator/controlling body/owner of resulting data please provide information on these other third parties and relative roles alongside SRUC:**

The grant is led by Liverpool School of Tropical Medicine, with SRUC, Roslin Institute and the Tsetse and Trypanosomiasis Research Institute, Tanga, Tanzania as co-PIs. The field work components of this project, including the cross-sectional survey of cattle, are the responsibility of TTRI. The survey will be coordinated and conducted by the TTRI veterinarian, who is experienced in working with cattle in similar situations. Harriet Auty, SRUC, is providing veterinary epidemiology expertise as needed to ensure the study is conducted in a robust epidemiological manner. The blood samples that are collected will be imported to the Roslin Institute for molecular analysis.

9) **What will any data/sample analyses collected be used for? Is it educational (student dissertation etc.), research (including publication/dissemination in academic circles or to support industrial development) or industry (used directly by third parties for management decisions, e.g. genetic evaluation data for industry):**

The samples will undergo molecular analysis at the Roslin Institute. The results from this will be fed back to farmers in the future, and will be discussed with farmers and other local livestock stakeholders such as the District Veterinary Officer, to help improve cattle disease control in the District. Within the BBSRC-funded project, the results will contribute to research conducted to elucidate transmission of human infective trypanosomes, in order to optimise control options to reduce human disease risk.

10) Potential benefits of the study:

The potential benefits include increased understanding of human African trypanosomiasis transmission and reservoirs, leading to optimised control and reduced human incidence of a disease which is fatal without treatment. In addition, the prevalence of cattle trypanosomiasis and East Coast Fever will be established. These are the most important livestock diseases in this area. Increased knowledge will ultimately lead to opportunities for improved control of livestock disease, also leading to impacts on food security and rural poverty.

11) What potential ethical issues does your study raise, and how have you addressed these? Has the study been approved by any other body (i.e. ethical review committee)?

The study as a whole has been approved by the Tanzania Wildlife Research Institute and the Tanzania Commission for Science and Technology. However, this approval process does not specifically cover the detailed study design. I am currently seeking information on the approval process that TTRI are required to go through (if any).

The potential ethical issues include two areas: (i) animal-related ethical issues to do with collection of diagnostic samples from animals, and (ii) human-related ethical issues to do with obtaining informed consent from livestock owners for sampling of animals and provision of information about those animals.

Other elements of the project, led by other partners, have received ethical approval from other organisations. For example, questionnaires will also be conducted with livestock owners, for which ethical approval has already been gained from the University of Edinburgh, as this part of the work is led by a PhD student at the Roslin Institute.

12) Will the staff/student require any training to carry out the project? (e.g. animal handling, data collection)

No training will be required. The study will be conducted by a Tanzanian veterinarian alongside a veterinary technician and local District Livestock officer

(government extension officer). They are all experienced in cattle handling in this type of environment, and in conducting this type of study in Tanzania. Epidemiological expertise on study design will be provided by Harriet Auty at SRUC who has experience running this sort of survey in Tanzania.

13) **Particularly for student projects, are you able to provide statistical guidance, or do you have sufficient knowledge to carry out any required analysis?**

NA – Within the consortium we have sufficient knowledge for all analysis.

14) **Are there any safety issues involved?**

Restraint of cattle for collection of samples can present a risk of injury to those involved in handling. However the restraint will be conducted by personnel experienced in handling cattle, under the supervision of a Tanzanian vet, to minimise risks. This work will be coordinated and conducted by TTRI, and not by SRUC.

15) **If you are a supervisor of a student, can you foresee any problem if the student is affiliated with an NGO/charity who looks after animal welfare/rights issues?**

NA.

COMMENTS AND RESPONSES FOR BORDERLINE FORM

AE Number:	NA
AE Title:	Cross sectional survey of *Trypanosoma spp.* prevalence in livestock in Tanzania.
RP Number:	NA
RP Title:	NA
Research Project Manager:	NA
Experimental Manager:	Harriet Auty
Date of start of trial:	July 2016
Date Submitted to AEC:	
Deadline for comments (initial comments due after 6 days, final comments due a maximum of 12 working days from date of circulation)	Initial deadline: Final deadline:
***Approved by:**	**Date:**
Vicky Sandilands (Chairperson)	25/05/16
Tony Waterhouse (Vice Chair)	26/5/16
Julia Chambers (NVS)	22/6/16
Colin Mason (NVS)	
Steven Thomson (NAP)	2/6/16
Laura Nicoll (NACWO)	13/06/2016
Mhairi Jack (NACWO)	
Hugh McClymont (NACWO)	
Ainsley Bagnall (NACWO)	
Ian Nevison (Statistician)	
Nicola Lambe (Scientist)	
Kenny Rutherford (Scientist)	30/5/16
Ann Bruce (External)	

**If you are content with the form as it stands, please date the box opposite your name.*

If you have a comment which is for consideration or information, but does not require to be addressed in order for you to accept the form, please also date the box above and insert your comment below.

If, for some reason, the form is unacceptable to you, insert the relevant reason below, but please precede the reason with the words "unacceptable because …"

If you do make a comment please ensure that you revisit this form in order to confirm you are happy with the scientists response.

Comments from committee (to be addressed by submitting scientist):

VS 25/05/16

I am satisfied that a Borderline form is suitable:

SRUC involvement is very minor

The blood sampling is diagnostic in nature – there are no treatments being applied

It will be interesting to know if, once diagnosed, something will be done to then help the situation – culling (of cattle, not people!)? Vaccination? Etc

HA Response: In the short term, the TTRI vet will provide advice on individual animals as required.. Longer term, the aim of the project is to develop improved control options for animal diseases, and to reduce the risk of human disease, ideally through integrated control that can address these human and livestock disease together. We anticipate conducting feedback workshops for farmers to learn about the results of these surveys, and the best options for control, in later years of the project.

AW 26/5/16

I agree – good test of our modified form?

I think outcome would be to approve Harriets/SRUC involvement if others on AEC have similar views. Clear that others, TTRI and Liverpool have more primary roles in the direct animal work.

JC 26/5/16

Please can you explain why you require this sample too: "In addition, a peripheral capillary sample will be collected from an ear vein, using a lancet and capillary tube"? Presumably there will be a lower prevalence in these samples compared with jugular blood in chronic T.*brucei* cases? Will there be an increased risk of bleeding, infection, insect attraction at blood sample sites in ECF- and *T. brucei*-infected cattle? If so, how will you prevent and monitor for that?

HA Response: In contrast to *T. congolense*, *T. brucei* localises in peripheral tissues rather than within the main blood circulation (Losos and Ibede 1972). Therefore ear capillary sampling may provide a more sensitive test of infection, particularly in chronically infected animals where the parasitaemia is low. It is widely used in trypanosomiasis surveys (for example Selby et al 2013, von Wissmann *et al.* 2014, Brownlow 2008). This study will routinely use jugular venopuncture because the small volume of blood acquired through ear capillary sampling (approx. 100ul) is insufficient for generating genetic material for further analysis. The aim for taking ear capillary samples from a subset of animals is to ensure our results are comparable with those from other studies.

Regarding the possibility of increased risk associated with jugular venopuncture or capillary sampling in infected animals, in general risks associated with venopuncture when carried out according to good practice are extremely low. The methods used involve such a small puncture that coagulation happens in 30-90 seconds and there is no significant bleeding or infection risk. Regarding ear capillary sampling specifically, in surveys of cattle populations where trypanosomiasis and East Coast fever are prevalent, although infections or haemorrhage were not assessed specifically, no issues were noted associated with increased risk of infection or haemorrhage in cattle, even in longitudinal studies where cattle/villages were revisited repeatedly (personal communication Beatrix von Wissmann and Andrew Brownlow).

To minimise the already extremely low risks, animals undergoing blood sampling by either method will not be released until it is clear that any bleeding has stopped. If for any reason an additional infection risk is suspected, antibiotic spray will be administered to the venipuncture site.

Losos G. and Ikede B. (1972) Review of pathology of diseases in domestic and laboratory animals caused by *Trypanosoma congolense, T. vivax, T. brucei, T. rhodesiense* and *T. gambiense*, *Veterinary Pathology* suppl 9

Selby, R. *et al.* (2013) Cattle movements and trypanosomes: restocking efforts and the spread of *Trypanosoma brucei rhodesiense* sleeping sickness in post-conflict Uganda. *Parasit. Vectors* 6, 281

von Wissmann, B. *et al.* (2014) Quantifying the Association between Bovine and Human Trypanosomiasis in Newly Affected Sleeping Sickness Areas of Uganda. *PLoS Negl. Trop. Dis.* 8,

Brownlow, A. (2007) Evaluation of a novel method for controlling bovine trypanosomiasis, PhD thesis, University of Edinburgh.

JC (22/6/16)

Many thanks. Hope all goes well.

VS 22/06/16 APPROVED