Modelling the spatial-temporal distribution of Tsetse (*Glossina pallidipes*) as a function of topography and vegetation greenness in the Zambezi Valley of Zimbabwe

Farai Matawa[1, 2], Amon Murwira[1], Fadzai M. Zengeya[1] and Peter M Atkinson[3,4,5]

1. University of Zimbabwe, Department of Geography and Environmental Science, P.O. Box MP167, Mt Pleasant, Harare, Zimbabwe
2. Scientific and Industrial Research and Development, Geoinformation and Remote Sensing Institute, 1574 Alpes Road, Hatcliffe Extension, Harare
3. Faculty of Science and Technology, Engineering Building, Lancaster University, Lancaster LA1 4YR, UK
4. School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK
5. Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Abstract

In this study, we developed a stable and temporally dynamic model for predicting tsetse (*Glossina pallidipes*) habitat distribution based on a remotely sensed Normalised Difference Vegetation Index (NDVI), an indicator of vegetation greenness, and topographic variables, specifically, elevation and topographic position index (TPI). We also investigated the effect of drainage networks on habitat suitability of tsetse as well as factors that may influence changes in area of suitable tsetse habitat. We used data on tsetse presence collected in North western Zimbabwe during 1998 to develop a habitat prediction model using Maxent (Training AUC=0.751, test AU=0.752). Results of the Maxent model showed that the probability of occurrence of *G. pallidipes* decreased as TPI increased while an increase in elevation beyond 800 m resulted in a decrease in the probability of occurrence. High probabilities (>50%) of occurrence of *G. pallidipes* were associated with NDVI between high 0.3 and 0.6. Based on the good predictive ability of the model, we fitted this model to environmental data of six different years, 1986, 1991, 1993, 2002, 2007 and 2008 to predict the spatial distribution of tsetse presence in those years and to quantify any trends or changes in the tsetse distribution, which may be a function of changes in suitable tsetse habitat. The results showed that the amount of suitable *G. pallidipes* habitat significantly decreased (r^2 0.799, p=0.007) for the period 1986 and 2008 due to the changes in the amount of vegetation cover as measured by NDVI over time in years. Using binary logistic regression, the probability of occurrence of suitable tsetse habitat decreased with increased distance from drainage lines. Overall, results of this study suggest that temporal changes in vegetation cover captured by using NDVI can aptly capture variations in habitat suitability of tsetse over time. Thus integration of remotely sensed data and other landscape variables enhances assessment of temporal changes in habitat suitability of tsetse which is crucial in the management and control of tsetse.

Keywords: NDVI, Elevation, Maxent, *Glossina pallidipes*, habitat

1.1 Introduction

The tsetse fly (Glossina spp.) is a vector that transmits the trypanosomes that are responsible for Human African Trypanosomiasis (HAT) in Humans, also known as sleeping sickness and African Animal Trypanosomiasis (AAT) in animals, which is often termed Nagana in cattle. The tsetse fly causes rural poverty across large areas of sub-Saharan Africa where the keeping of livestock is curtailed or prevented (Holmes, 2013, Matawa *et al.*, 2013). It is therefore important to understand the spatial-temporal dynamics of the tsetse flies in order to effectively apply vector control and eradication measures in order to improve rural livelihoods. The distribution of tsetse is often linked to specific habitat types, particularly those places with vegetation cover including thickets and riverine woodlands that provide ample shade and reduce the chances of dehydration (Adam *et al.*, 2012, Batchelor *et al.*, 2009, Odulaja and Mohamed-Ahmed, 2001, Van den Bossche *et al.*, 2010). Such habitats are also home to wildlife species that provide the requisite blood meals for the tsetse fly (Ducheyne *et al.*, 2009, Van den Bossche *et al.*, 2010). Thus, any landscape change that results in thicket reduction could affect not only the wildlife species but also affect the tsetse population both directly and indirectly (Kitron *et al.*, 1996, Munang'andu *et al.*, 2012). We therefore assert that characterisation of landscape changes is critical to understanding changes in the tsetse population and its distribution. Such characterisation also has potential to provide insights into the temporal and spatial dynamics of AAT in domestic animals and HAT in humans within ecosystems that are home to the tsetse fly.

Although an understanding of the spatial dynamics of key ecosystems is critical in characterising the dynamics of Trypanosomiasis, studies on ecosystem change and its effect on tsetse habitat dynamics have remained limited. Of the few studies on ecosystem change, the focus has mainly been on agricultural and human settlement expansion following the suppression of tsetse (Baudron *et al.*, 2010, Sibanda and Murwira, 2012a) and the consequent wildlife habitat changes. Understanding ecosystem change in relation to tsetse habitat could provide improved insights into how these changes alter the interactions between the host, vector and parasite (DeVisser *et al.*, 2010, Van den Bossche *et al.*, 2010). However, in order to track fine scale environmental changes, as well as, link these changes to tsetse fly presence or abundance there is need for the development of spatially explicit models at a fine spatial resolution (Rogers *et al.*, 1996) that incorporate dynamic variables that are able to capture changes in landscape condition.

The distribution of tsetse has been widely linked to vegetation cover as it influences micro-climate and availability of hosts (Cecchi *et al.*, 2008, DeVisser *et al.*, 2010, Hay *et al.*, 1997, Welburn *et al.*, 2006). Vegetation cover inherently changes over time and hence could be a useful dynamic variable that can be included in habitat suitability models. However, traditional approaches of quantifying vegetation cover have often been tedious, time consuming and limited to small areas. To this end, objective measures of quantifying vegetation cover over large spatial extents are thus important.

The advent of remotely sensed data has allowed objective measures of vegetation cover to be developed. For example, remotely sensed indices such as Ratio vegetation index (RVI), the Transformed vegetation index (TVI) and the Normalised Difference Vegetation Index (NDVI) have been developed to estimate vegetation cover across landscapes. Among these indices, NDVI has been widely used for characterizing vegetation cover, vegetation biomass and vegetation greenness (DeVisser *et al.*, 2010, Dicko *et al.*, 2014, Robinson *et al.*, 1997, Rogers *et al.*, 2000). For example, NDVI in combination with temperature and rainfall were used to explain the distribution of tsetse flies in West Africa based on the discriminant analysis approach (Rogers *et al.*, 1996). Although these studies have provided insights into factors influencing the distribution of tsetse, the studies failed to take into account temporal variation in tsetse habitat.

Furthermore, the remotely sensed data used in these studies particularly NDVI was derived from low resolution satellite data which tend to over-generalise tsetse habitat. It is well known that tsetse populations can be maintained in small patches of suitable habitat particularly micro-habitats provided by land cover types that contain woody vegetation (DeVisser *et al.*, 2010). Thus habitat suitability models developed using low resolution NDVI data derived from 250 m MODIS and 1 km NOAA-AVHRR sensors may fail to capture patches of suitable habitat smaller than 250 m spatial resolution (DeVisser *et al.*, 2010). Furthermore, use of low spatial resolution imagery may compromise the results of epidemiological analyses (Atkinson and Graham, 2006). In this regard, inclusion of remotely sensed estimates of vegetation cover at a fine resolution is imperative in enhancing the accuracy and usefulness of tsetse distribution models in tsetse eradication campaigns.

In this study, our main objective was to assess temporal changes in *G. pallidipes* habitat based on a habitat model developed using dynamic and stable environmental variables. We hypothesised that ecosystem changes resulting from changes in landcover reduce the amount of suitable tsetse habitat. Specifically, we tested whether *G. pallidipes* habitat can be predicted based on three variables namely 30 m resolution Landsat TM based Normalised Difference Vegetation Index (NDVI) (temporally dynamic variable) as well as elevation and Topographic Position Index (TPI) (temporally stable variables). We then tested the ability of the model to predict tsetse suitable habitat for 1986, 1991, 1993, 2002, 2007 and 2008 in order to characterise the spatial dynamics of tsetse habitat over time. We also tested whether suitable tsetse habitat varied temporally due to reduction in vegetation cover. We explained the relationship between spatial temporal variation in suitable habitat and rainfall as well as burnt area and assessed whether there are net gains or losses in suitable habitat between successive years.

We considered topographic variables such as elevation and TPI due to the fact that Tsetse is mostly found in low-lying areas as they are associated with high temperatures (DeVisser *et al.*, 2010, Matawa *et al.*, 2013, Terblanche *et al.*, 2008). TPI measures slope position and landform category i.e. identifies hilltops, ridges, valleys and flat areas (Pittiglio *et al.*, 2012). However, elevation and TPI may fail to capture the spatial-temporal dynamics in tsetse fly occurrence as they are largely temporally stable. Thus their integration with remotely sensed vegetation cover could provide a spatially and temporally dynamic model that can allow modelling of changes in tsetse suitable habitat over time.

1.2 Materials and Methods

1.2.1 Study Area

The study area is located in north western Zimbabwe at 16˚ south and 29˚ east (Figure 1). The study was conducted in an area straddling protected areas (including safari areas) and settled areas comprising large and small scale farming areas and the communal lands of the Zambezi Valley. Communal lands are areas characterised by community land ownership and are subdivided into administrative units called wards (Sibanda and Murwira, 2012a).

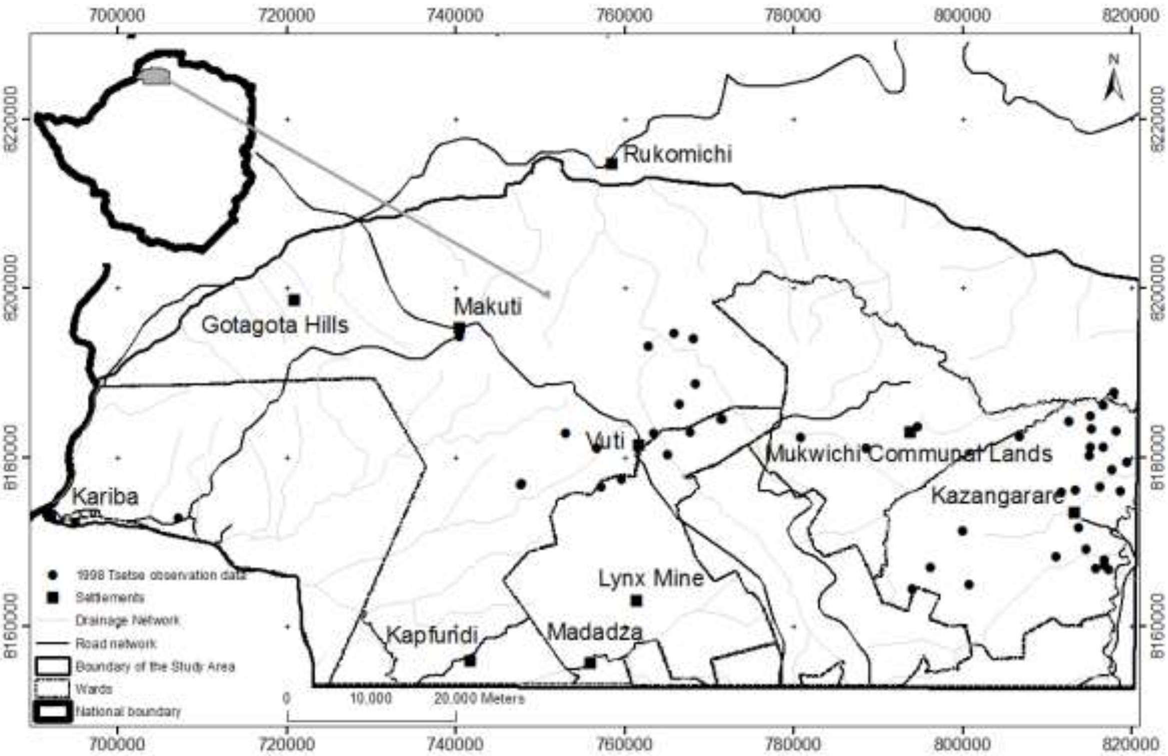

Figure 1: Location of the study area in Zimbabwe

The area has a dry tropical climate, characterised by low and variable annual rainfall averaging between 450 and 650 mm per year and a mean annual temperature of 25°C (Baudron *et al.*, 2010, Sibanda and Murwira, 2012a). The rainfall patterns based on mean monthly precipitation calculated using data recorded at the three closest whether stations namely Karoi, Makuti and Rekomitje (Rukomichi) show that the 1985/1986 rainfall season had higher rainfall as compared to all the other rainfall seasons under consideration (Figure 2**Error! Reference source not found.**). The area has two clearly defined seasons: a wet season from December to March and a long dry season from April to November (Baudron *et al.*, 2010). The climatic conditions, thus, make the study area a suitable habitat for tsetse. The natural vegetation is mainly deciduous dry savannah, that includes *Colophospermum mopane* (Baudron *et al.*, 2010, Sibanda and Murwira, 2012a), *Combretum* woodlands and riparian vegetation. The elevation of the study area ranges from 340 m to 1400 m (SRTM-DEM). Areas below 1100 m are climatically suitable for tsetse (Pender *et al.*, 1997).

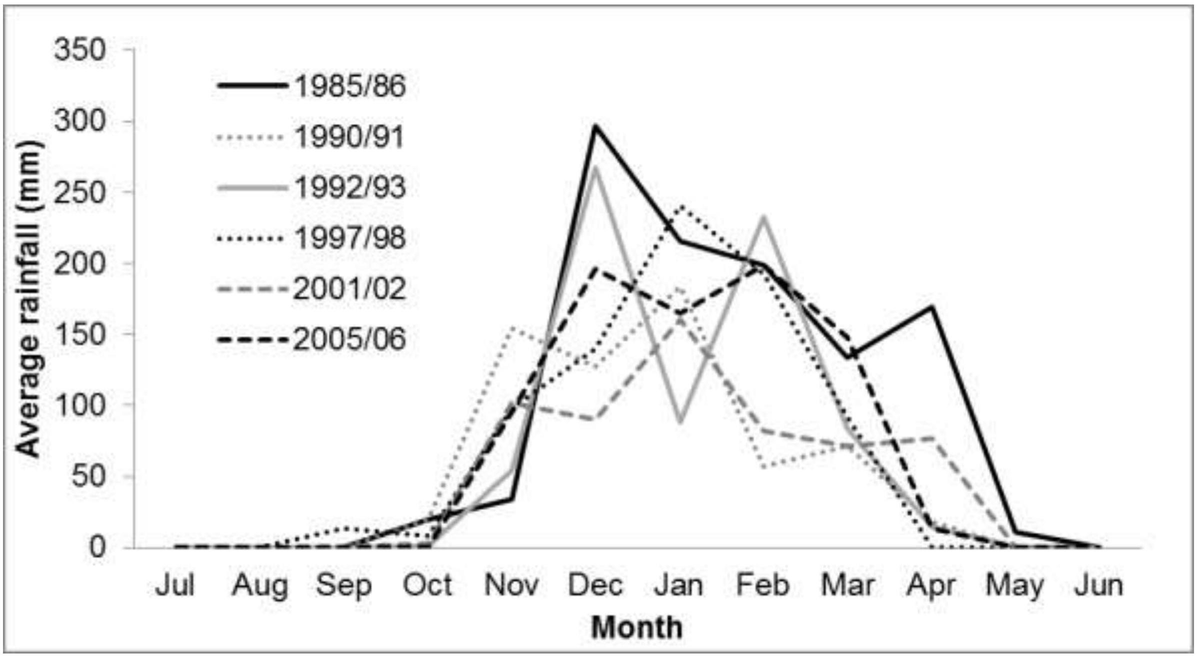

Figure 2: Rainfall patterns based on Karoi, Makuti and Rukomichi weather stations (Source: Meteorological Services Department).

The major economic activity is dryland farming of cotton (*Gossypium hirsutum*), maize (*Zea mays*) and sorghum (*Sorghum bicolor*) (Baudron *et al.*, 2010) as well as tobacco. There have been initiatives by government since 1960 to eradicate tsetse in the study region and this has resulted in the tsetse front progressively receding towards the Zambezi River (Shereni, 1990).

1.2.2 Species Occurrence Data

Data on tsetse occurrence were extracted from tsetse fly trapping records for the period 1994 to 2012. We used the 1998 dataset for training the model because it had a better spread and more presence records (50). The tsetse fly trap records were collected by the Zimbabwe Department of Veterinary Services and Livestock Production, Tsetse Control Division in Harare. The tsetse distribution data were however collected by marking the tsetse sightings on 1:250 000 scale maps. In order to allow the data to be integrated with other spatial data sets we first scanned the maps and georeferenced them in a GIS (RMSE=0.000033). Next, we digitized the tsetse sighting locations (Figure 1).

1.2.3 Environmental variables

We downloaded cloud-free (less than 10% cloud) 30m spatial resolution Landsat TM satellite sensor data made available at the USGS EROS Data Centre (http://lpdaac.usgs.gov/) in order to estimate vegetation greenness. Satellite sensor data collected were for the period April to early-July (day 110 to day 199) for the years 1986, 1991, 1993, 1998, 2002, 2007 and 2008. We focused on the period from end-April to early-July (post-harvest period) as all trees in the study area are still in full leaf while grass and crops would be in the senesce stages (Sibanda and Murwira, 2012b) thereby making it easier to explain the impact of land use/ landcover change. The Landsat TM and ETM data were already georeferenced to the Universal Transverse Mercator (UTM) Zone 35 South based on the WGS84 spheroid. However, we checked for the accuracy of the georeferencing based on 20 ground control points (i.e. river intersections) from georeferenced 1:50 000 topographic maps of the study area. Vegetation greenness was estimated using the Normalised Difference Vegetation Index (NDVI) as follows:

$$NDVI = \frac{NIR - R}{NIR + R}$$

where *NIR* is the reflectance in the near infrared wavelength while *R* is reflectance in the red wavelength of the electromagnetic spectrum. We used NDVI as it is a good estimator of vegetation greenness, vegetation cover and vegetation biomass (Huete *et al.*, 2002). We calculated average NDVI based on available Landsat TM and ETM imagery between day 110 and day 199 of each year. We selected years with at least two or more images for the analysis. We masked out clouds to reduce their influence on the average NDVI values and outcome of the model. We then calculated the average NDVI for each year based on the period end-April to early-July.

Next, we used the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) at a spatial resolution of 90 meters (*www.usgs.gov*) and then resampled to 30m spatial resolution to estimate topographical variables, i.e. elevation and topographic position index in a GIS.

We also acquired and processed readily available MODIS burnt area data for 2002, 2007 and 2008 in order to measure the burnt area coinciding with the Landsat TM data (modis-fire.umd.edu). The burnt area data for 1986, 1991, 1993, and 1998 was not readily available. Therefore we could not fit a regression model between burnt area and suitable habitat. We used this data to explain the link between variations in burnt area and the fluctuations in area of suitable tsetse habitat using a graphical plot.

We acquired rainfall data for Karoi, Makuti and Rukomichi weather stations for the 1985/ 86, 1990/91, 1992/1993, 1997/98 and 2001/2002 rainfall seasons to explain the suitable habitat for 1986, 1991, 1993, 1998 and 2002 respectively and calculated an average seasonal total from the 3 stations. The rainfall data for 2006/07 and 2007/08 was not readily available to be used to explain the suitable habitat for 2007 and 2008 respectively. Therefore out of the 7 years under consideration we only analysed the relationship between suitable habitat and rainfall for only five years.

1.2.4 Modelling tsetse habitat using the Maximum Entropy method

We used the Maximum entropy (Maxent) modelling approach (Phillips *et al.*, 2006, Phillips and Dudik, 2004) to predict the spatial distribution of tsetse in the study area as a function of elevation, topographic position index (TPI) and NDVI. Maxent utilises presence only data to model habitat suitability as a function of environmental variables. In this study, we used presence only data because tsetse presence data are generally more meaningful than absence data as all known traps have a very low efficiency with respect to trapping rates and therefore there are chances of generating false absence data (Dicko *et al.*, 2014, Rogers *et al.*, 1996). We treated tsetse trap records as presence only data that could be used to model tsetse habitat suitability as a function of NDVI, TPI and elevation.

For the modelling process, tsetse occurrence data (n =50) (Figure 1) for the year 1998 were randomly partitioned into a 70% training subsample and a 30% test subsample (Matawa *et al.*, 2012). We used the 1998 tsetse location data to build the initial model because it had more data points than the other years in the Tsetse Control Division database as well are more than one image for the post-harvest period. In order to evaluate the accuracy of the model we used the area under curve (AUC) of the receiver operating characteristics (ROC) (Phillips *et al.*, 2006, Phillips and Dudik, 2004). AUC values range from 0 to 1 where values between 0 and 0.5

reflect that the model fails to establish habitat suitability for the tsetse while values closer to 1 indicate that the model successfully establishes the suitable habitat. In fact, AUC values between 0.7 and 0.80 are classified as average in terms of model accuracy while AUC values between 0.6 and 0.70 are classified as poor (Parolo *et al.*, 2008).

The Maxent model determined using the 1998 data was then used to predict tsetse habitat suitability in 1986, 1991, 1993, 2002, 2007 and 2008 using appropriate covariate images. We then converted the probability maps into binary maps (i.e. suitable (1) and unsuitable (0)) using the 'equal training sensitivity and specificity' threshold rule in Maxent (Phillips *et al.*, 2006).

1.2.5 Assessment of the spatial temporal dynamics of G. pallidipes habitat

In order to understand the variations in suitable and unsuitable habitat between land cover/ use types, we extracted suitable and unsuitable areas within communal lands and protected areas using overlay analysis in the Integrated Land and Water Information System (ILWIS) geographic information system software (www.52North.org). The same procedure was also followed for riverine and non-riverine areas. We then calculated the area of suitable and unsuitable tsetse habitat that fell within the land cover/ use types using the area calculation function in ILWIS. The riparian/ riverine forest was delineated by creating a 500 m buffer along the stream network similar to the one used by (Guerrini *et al.*, 2008) whilst the non-riparian forest was the area beyond 500 m from river courses. We compared the proportion of suitable riparian habitat to the proportion of suitable non-riparian habitat in the communal lands using the Z-score test in the R software (https://cran.r-project.org/package). The test for proportion (Z-score test) is formulated as follows:

$$z = \frac{(\overline{p}_1 - \overline{p}_2) - 0}{\sqrt{\overline{p}(1 - \overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$\overline{p}$ is the sample proportion, and n is the sample population (Agresti and Coull, 1998).

We also tested whether or not the proportion of suitable habitat is significantly different to the proportion of unsuitable habitat in the communal lands where there is dense human activity. We used the Z-score test to test for differences between proportions.

In order to determine whether there was a general trend over time in habitat suitability we related the area of suitable habitat with time in years to confirm the trend of decrease in suitable habitat over time using the exponential model. We also calculated the net loss and gains in tsetse habitat for the study period by accounting for changes in area of suitable habitat between the current year and the previous (base) year, i.e., between 1986 and 1991, 1991 and 1993, 1993 and 1998, 2002 and 2007 and 2007 and 2008 as well as between 1986 and 2008.

1.2.6 Influence of drainage network on habitat suitability

We used binary logistic regression (Pearce and Ferrier, 2000) to investigate the relationship between the drainage network and the distribution of suitable and unsuitable of tsetse habitat of 1986, 1991, 1993, 1998, 2002, 2007 and 2008. We generated 1000 random points for the whole study area using the random points option in QGIS (www.qgis.org). We then used these points to extract binary data from the model outputs of all the years under consideration and distance from the main drainage network of the study area using the overlay function in a GIS. The distance from the main drainage network was calculated based on the Euclidian distance from the drainage network in ILWIS (Matawa et al., 2012). We then related the binary data for each year with distance from the drainage network. Binary logistic regression is formulated as follows:

$$P = \frac{\exp(\beta_0 + \beta_1 * X_1)}{(1 + \exp(\beta_0 + \beta_1 * X_1)}$$

P is the probability of the outcome occurring β_0 is the constant, β_1 is the gradient and X_1 is the independent variable of the equation. Model performance was evaluated by considering the area under the Receiver Operator Characteristic curve (ROC).

1.2.7 Factors explaining the changes in habitat suitability

To explain the fluctuations in the changes in suitable habitat across time, we analysed the relationship between suitable habitat and average seasonal rainfall based on Kariba, Makuti and Rukomichi weather stations using linear regression in the Statistical Package for Social Scientists (SPSS) from 1998 to 2002. This was based on the assumption that fluctuations in suitable tsetse habitat derived from NDVI data can be explained by rainfall variability. Prior to analysis we tested whether the data followed a normal distribution using Kolmogorov-Smirnov test. Results showed that data did not deviate from a normal distribution (p=0.2). In addition, we calculated and compared the proportions of burnt area for 2002, 2007 and 2008 as well as generating plots of the burnt area in order to allow visual comparison with the modelled suitable tsetse habitat. .

1.3 Results

1.3.1 Species Distribution Model

The AUC values obtained for the 1998 model as a function of elevation, TPI and NDVI are greater than 0.5 showing a significant departure from randomness (Table 1).

Table 1: AUC values for the individual variables and the overall Maxent model

Variable	AUC-value	AUC for training data (70%)	AUC for test data (30%)
Elevation	0.663		
Topographic position index(TPI)	0.659		
NDVI	0.739		
Overall Maxent model		0.751	0.752

Based on the model results, the probability of occurrence of *G. pallidipes* decreases with an increase in TPI as high TPI is associated with elevated areas such as hilltops and low TPI values are associated with valleys (Figure 3a). Figure 3b shows that the probability of *G. pallidipes* decreases sharply when elevation exceeds 1100 m. The probabilities of 50% and above are associated with NDVI values of between 0.4 and 0.6 (Figure 3c).

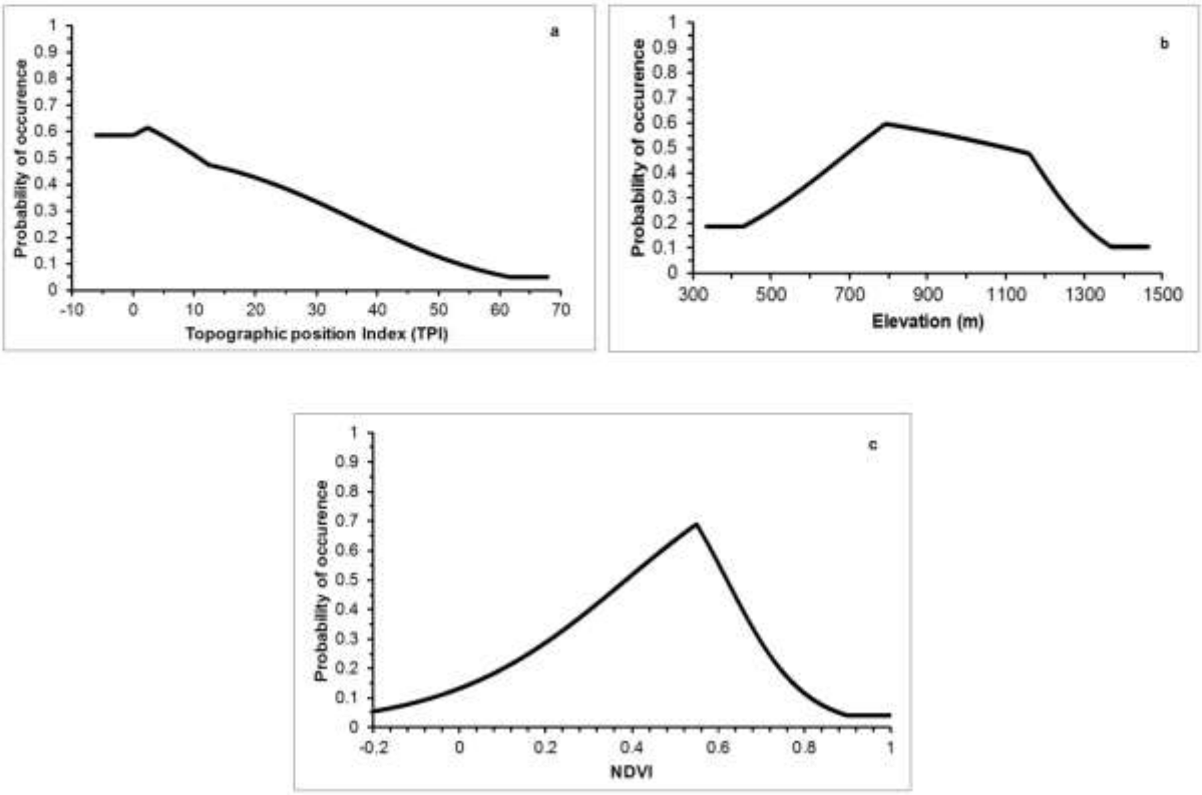

Figure 3: Relationship between (a) TPI, (b) Elevation and (c) NDVI and probability of presence of G. pallidipes.

Using the 1998 model to predict tsetse habitat suitability for the period 1986, 1991, 1993, 2002, 2007 and 2008 it can be observed that there are marked spatial shifts in the suitable habitat for *G. pallidipes* from 1986 to 2008 in the communal lands (Figure 4a-g). We observe that the smallest patch of suitable habitat identified by our model is 900 m^2. Figure 4 also illustrates that the suitable habitat is also concentrated along riverine areas, for example the Mushangizhi and Mukwichi Rivers.

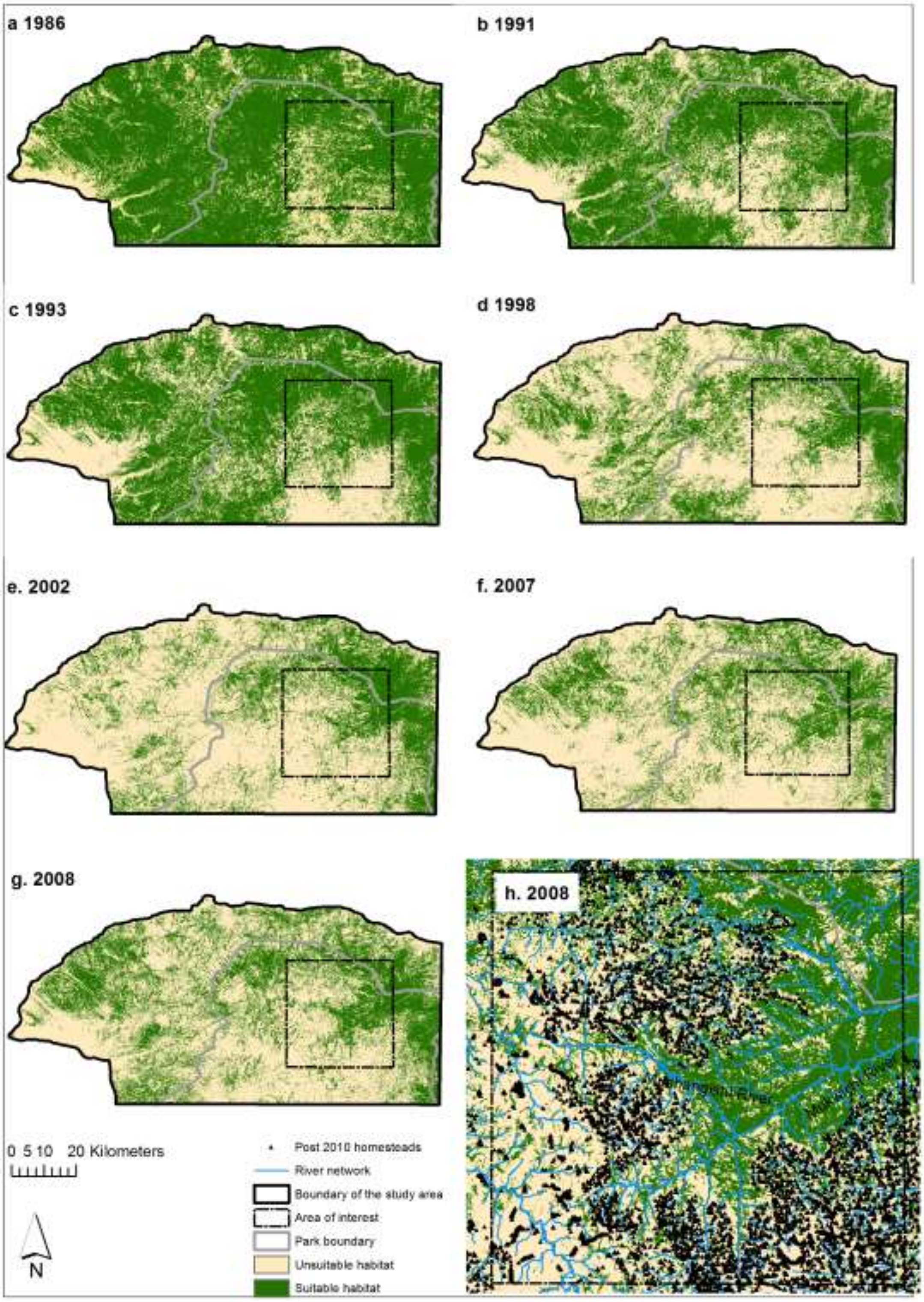

Figure 4: Spatial-temporal variation in the distribution of suitable and unsuitable Glossina pallidipes habitat from a) 1986, b) 1991, c) 1993, d) 1998, e) 2002, f) 2007 and g) 2008 based on changes in vegetation cover. The area bounded by the black dashed box illustrates a settled area where human activity is intense and shows changes in suitable habitat between 1986 and 2008 and is zoomed in 4(h).

The location of post 2010 homesteads in the study area is coinciding mostly with unsuitable *G. pallidipes* habitat (Figure 4h). This is also the area where agricultural activity is intense in the study area. Some areas that were suitable in 1986 were now unsuitable habitat in 2008 (Figure 4).

1.3.2 Assessment of the spatial temporal dynamics of G. pallidipes habitat in the communal lands

Tsetse habitat receded between 1986 and 2002, and then it increased slightly between 2002 and 2008 in the communal lands. However, the proportion of suitable habitat modelled for 2008 is significantly lower than the 1986 proportion of suitable habitat ($p = 0.00001$). For the period 1986 to 1993 the proportion of suitable habitat was significantly higher ($p<0.05$) than the proportion of unsuitable habitat whilst the period 1998 to 2008 the proportion of unsuitable habitat was significantly higher ($p<0.05$) than the proportion of suitable habitat between 1998 and 2008 (Table 2).

Table 2: Comparison of the proportions of suitable habitat and unsuitable habitat in the communal lands characterised by dense human activity using the Z-test at the 95% Confidence Interval. (The values in brackets in the second and third columns are proportions)

Year	Suitable Habitat (ha)	Unsuitable Habitat (ha)	Standard Error (S.E.)	Lower Bound	Upper Bound	Z-score	p-value
1986	213694	83936	0.000008	0.7179686	0.7180010	336.365	0.00001
	(0.7179854)	(0.2820145)					
1991	165829	131801	0.000009	0.5571468	0.5571825	88.209	0.00001
	(0.5571649)	(0.4428351)					
1993	164281	133349	0.000009	0.5519463	0.5519820	80.1834	0.00001
	(0.5519638)	(0.4480362)					
1998	122575	175055	0.000902	0.4118185	0.4118539	-136.04	0.00001
	(0.4118368)	(0.5881632)					
2002	104227	193403	0.000874	0.3501738	0.3502081	-231.17	0.00001
	(0.3501898)	(0.6498102)					
2007	107969	189661	0.000881	0.3627446	0.3627791	-211.77	0.00001
	(0.3627625)	(0.6372375)					
2008	116858	180772	0.000895	0.3926112	0.3926463	-165.68	0.00001
	(0.3926284)	(0.6073716)					

The model shows that, for the whole study area, there was a net decrease of suitable habitat by 84,562 ha between 1986 and 1991; 109,142 ha between 1991 and 1993; 48,417 ha between 1993 and 1998 as well as a net gain of 12,467 ha between 2002 and 2007 and 39,700 ha between 2007 and 2008. Overall, the model shows a net loss of 199,955 ha between 1986 and 2008.

We found a significant negative exponential relationship between modelled suitable tsetse habitat and time in years (r^2 =0.799, p=0.007) (Figure 5).

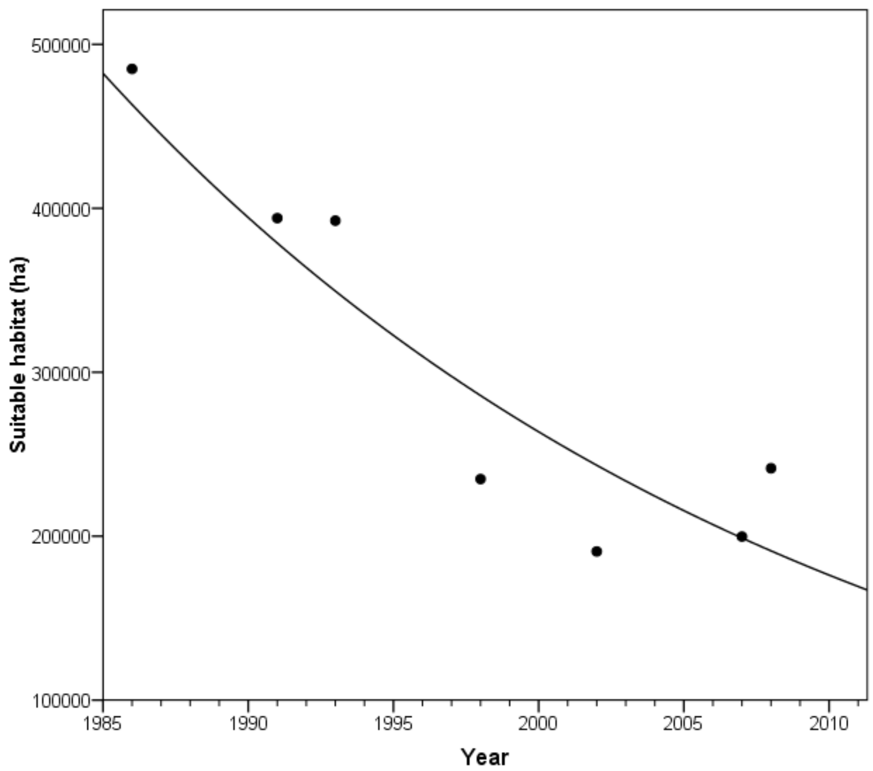

Figure 5: Relationship between suitable habitat and time in years

1.3.3 Influence of drainage network on habitat suitability

We observed that the proportion of suitable riverine habitat is relatively higher than the proportion of suitable non-riverine habitat in the communal lands (Table 3).

Table 3: Comparison of proportions of suitable riverine and suitable non-riverine habitat in the communal lands using the Z-test at 95% Confidence Interval.

Year	Suitable riverine habitat	Suitable non-riverine habitat	Z-score	P-value
1986	0.760	0.702	33.3941	0.00001
1991	0.612	0.534	40.4002	0.00001
1993	0.606	0.529	40.0483	0.00001
1998	0.487	0.385	52.9044	0.00001
2002	0.558	0.318	125.484	0.00001
2007	0.682	0.334	179.3278	0.00001
2008	0.563	0.360	104.9968	0.00001

We observed that the probability of occurrence of suitable habitat decreases with an increase in distance from the drainage network in the study area (Table 4). For 1998 and 2002 the relationship is statistically significant ($p<0.05$) (Table 4). Although for 1986, 1991, 1993, 2007 and 2008 the relationship is not statistically significant, all models show a trend of decrease of the probability of occurrence of suitable with an increase in the distance from the drainage network except for 1991. All models, except for the 1991 model, performed better than random and have AUC values between 0.5 and 0.6 (Table 4).

Table 4: Relationship between suitable habitat and distance from the drainage network. The standard error is shown in brackets.

Year	**1986**	**1991**	**1993**	**1998**	**2002**	**2007**	**2008**
Intercept	1.067** (0.127)	0.329** (0.111)	0.388** (0.112)	-0.407** (0.114)	-0.548** (0.120)	-0.726** (0.122)	-0.450** (0.115)
p-value	0.0000	0.0032	0.0005	0.0004	0.00001	0.00000	0.0001
Stream Distance	-0.00003 (0.00012)	0.00002 (0.0001)	-0.00003 (0.0001)	-0.00024** (0.00011)	-0.00046** (0.00012)	-0.00015 (0.00012)	-0.00011 (0.00011)
p-value	0.78842	0.85287	0.80537	0.02208	0.00009	0.18264	0.28924
AUC	0.513	0.494	0.513	0.538	0.572	0.524	0.530

**Significant at 95% confidence interval

1.3.4 Factors explaining the changes in habitat suitability

Our results show that the seasonal variation of rainfall can positively explain fluctuations in NDVI derived suitable habitat change (r^2=0.977, r^2 adjusted =0.972, p=0.000192). In addition, the results show that the proportion of burnt area of 2002 is significantly higher than the proportion of burnt area in 2007 (z=65.0156, p=0.00001) and significantly higher than the proportion of burnt area in 2008 (z=111.09, p=0.0000). The proportion of burnt area for 2007 is significantly higher than the proportion of burnt area in 2008 (z=46.7473, p=0.00001). In addition, Figure 6a shows that the amount of suitable habitat was increasing between 2002 and 2008 as the amount of burnt area was decreasing.

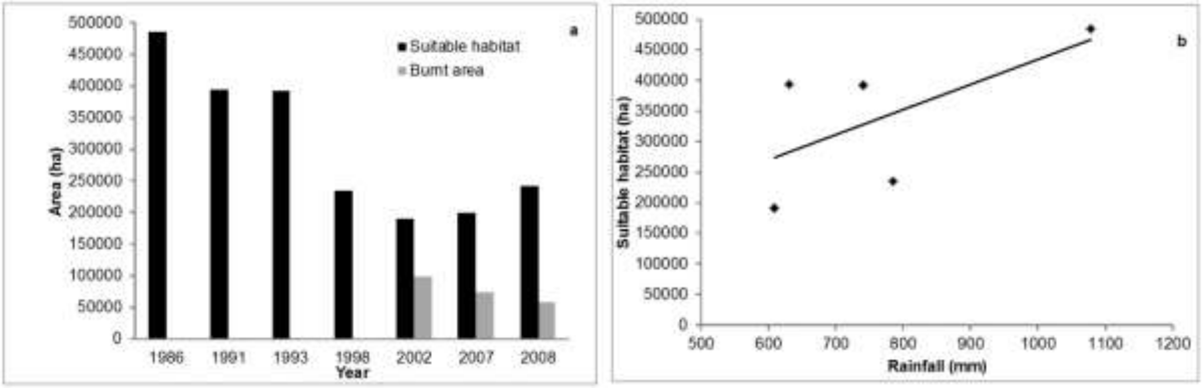

Figure 6: Relationship between (a) burnt area and amount of suitable habitat and (b) rainfall and amount of suitable habitat

1.4 Discussion

Results of this study indicate that spatial and temporal variability in vegetation cover affect the distribution of suitable tsetse habitat. The results indicate that changes in tsetse habitat are not uniform and unidirectional. Significant spatial changes (contraction and expansion) in suitable tsetse habitat were noted throughout the study period (1986-2008). There was however a general decline in suitable habitat of tsetse between 1986 and 2008. Our results are consistent with our hypothesis that changes in landcover which lead to ecosystem changes reduce the amount of suitable tsetse habitat. This study uses data covering seven years spanning a period of 12 years to understand the spatial and temporal dynamics of *G. pallidipes* habitat in response to landcover change. Although other studies focused on the fragmentation of the riparian habitat and its effect on tsetse distribution (Guerrini *et al.*, 2008), the data used was not multi-temporal. We therefore assert that inclusion of both stable and dynamic variables in spatially explicit habitat models improves the detection of habitat suitability changes in response to changing environment.

Results of this study indicate that NDVI in addition to topographical variables such as elevation and topographic position index can successfully predict changes in *G. pallidipes* habitat over time. The combination of these variables enabled a dynamic approach to modelling changes in habitat suitability of tsetse in response to changes in habitat condition. NDVI provides the dynamic part of the model while the TPI and elevation provide the stable part of the model. Our results are consistent with other findings in the Zambezi valley which showed that NDVI and elevation significantly predict tsetse habitat (Matawa *et al.*, 2013). However, unlike previous studies, this study focused on producing a temporal dynamic model for demonstrating how changes in landcover and associated ecosystem can trigger changes in habitat suitability of *G. pallidipes* at 30m spatial resolution.

The response curves for *G. pallidipes* probabilities are consistent with results of earlier studies. For example, DeVisser *et al.*, 2010, Matawa *et al.*, 2013 and Terblanche *et al.*, 2008 found that the tsetse is mostly found in low-lying areas as they are associated with high temperatures. The importance of vegetation cover on tsetse distribution due to its provision of shade and its influence on availability of hosts has been alluded to (Cecchi *et al.*, 2008, DeVisser *et al.*, 2010, Hay *et al.*, 1997, Welburn *et al.*, 2006). To the best of our knowledge TPI has not been applied to model tsetse distribution. TPI helps determine whether or not the species prefer valleys to hilltops as suitable habitat. Our study was able to demonstrate that TPI can explain tsetse habitat preference as well as that *G. pallidipes* prefers valleys to hilltops.

Our results show that unsuitable *G. pallidipes* habitat is coinciding with areas were human activity is intense as represented by homesteads digitised from high resolution Google and Bing based satellite imagery of post 2010 (Figure 4). This shows that the settlement of people and subsequent expansion of agriculture induced landcover changes and fragmentation of woodland areas (Sibanda and Murwira, 2012a). The loss of landcover in the post suppression period reduces the chance of re-invasion by tsetse flies as the ecological factors that support tsetse survival particularly presence of tree canopy cover were altered. Thus landuse, particularly intensification of agriculture, has a negative impact on the spatial distribution of *G. pallidipes*. This is consistent with Van den Bossche, 2010 who observed that intensification of human activity reduced the amount of suitable habitat for tsetse in Zambia. Population growth occurring in rural areas, may lead to reduction of tsetse habitat and a reduction in sleeping sickness risk due to alteration of landcover (Welburn *et al.*, 2006). Thus human activities such as practising arable agriculture can induce landcover changes that can reduce or eliminate tsetse habitat.

Changes in suitable habitat could be explained by variations in rainfall from year to year and fire scars (Figure 6) that have a direct impact on the amount of vegetation cover as shown by trends in relationship between rainfall and the trends in the proportion of burnt area in the study area. For example, the smallest area of suitable habitat was estimated in 2002 and the annual total rainfall was low and the monthly rainfall was erratic based on data from the 3 nearest weather stations (Figure 2). Vegetation cover as measured by NDVI is dependent on amount of rainfall as much as it is dependent on changes in landuse patterns from time to time. The suitable habitat in the communal lands where human activity is intense is mostly suitable along the riverine areas (Figure 4) and this was also confirmed using binary logistic regression. Unsuitable habitat is related to cultivation and grassland classes (FAO, 1996). This suggests that alteration of vegetation cover due to cultivation and other human activities can reduce the suitable habitat of *G. pallidipes.*

We were able to demonstrate that landcover change in the study area, particularly in the communal lands, has impacted more on the non-riverine habitat. The suitable habitat is mostly around riverine areas and valleys. The vegetation cover of these areas is less disturbed as compared to the non-riverine areas. This could be as a result of the location of agricultural fields away from major river channels. The difference between the proportion of suitable riverine habitat and the proportion of suitable non-riverine habitat in the communal lands in the study area can be explained by settlement and associated human activities concentrated on the plateau area avoiding rivers and valleys. Thus landuse change and associated landcover change has altered the habitat of tsetse flies in the post-suppression period such that it may be difficult for tsetse flies to re-establish critical populations in the settled parts of the study area.

This study differs from other studies in evaluating how landcover change over time influences the amount of suitable habitat available to *G. pallidipes* thereby developing an understanding of the tsetse habitat dynamics and the utility of the spatial temporal approach to characterising tsetse distribution. We were able to trace the changes in tsetse distribution from the 1980s, i.e. the early days of human immigration (Baudron *et al.*, 2010) to the post 2000 period. Although the Landsat TM and ETM data we used in this study suffers from low temporal fidelity compared to other sensors e.g. MODIS it offers a better spatial resolution which may improve the identification of isolated suitable tsetse habitats as small as 900m^2. This is similar to the smallest patch of suitable habitat that we identified in this study. This helps in enhancing the monitoring of tsetse prevalence, planning tsetse eradication and monitoring the effectiveness of tsetse eradication programmes. Overall, the model developed in this study allows environmental changes to be linked with changes in tsetse fly occurrence.

Conclusion

We conclude that ecosystem changes induced by landcover changes as measured by the remotely sensed normalised difference vegetation index (NDVI) can be used to track changes in tsetse habitat change on a spatial-temporal scale. The spatial heterogeneity in landcover as measured by remotely sensed NDVI can explain the spatial temporal dynamics of tsetse habitat. We were able to track the expansion and contraction of tsetse as NDVI varied with each rainfall season. Therefore landcover change has a significant impact on change in suitable tsetse habitat.

We conclude that our model can be used to track spatial-temporal changes in suitable tsetse habitat. This shows that *G. pallidipes* habitat varies from place to place and time to time due to changes in the amount of vegetation cover as measured by the normalized difference vegetation index (NDVI). We also conclude that loss of vegetation cover has reduced the amount of suitable *G. pallidipes* habitat in the Zambezi Valley of Zimbabwe.

References

Adam, Y., Marcotty, T., Cecchi, G., Mahama, C. I., Solano, P., Bengaly, Z. & Van Den Bossche, P. (2012) Bovine trypanosomosis in the Upper West Region of Ghana: Entomological, parasitological and serological cross-sectional surveys. *Research in Veterinary Science,* 92**,** 462-468.

Agresti, A. & Coull, B. A. (1998) Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician,* 52**,** 119-126.

Atkinson, P. M. & Graham, A. J. (2006) Issues of scale and uncertainty in the global remote sensing of disease. . *Advances in Parasitology,* 62**,** 79-118.

Batchelor, N. A., Atkinson, P. M., Gething, P. W., Picozzi, K., Févre, E. M., Kakembo, A. S. L. & Welburn, S. C. (2009) Spatial Predictions of Rhodesian Human African Trypanosomiasis (Sleeping Sickness) Prevalence in Kaberamaido and Dokolo, Two Newly Affected Districts of Uganda. *PLoS Neglected Tropical Disease,* 3.

Baudron, F., Corbeels, M., Andersson, J. A., Giller, K. E. & Sibanda, M. (2010) Delineating the drivers of waning wildlife habitat: the predominance of cotton farming on the fringe of protected areas in the Mid Zambezi Valley, Zimbabwe. *Biological Conservation,* 144**,** 1481–1493.

Cecchi, G., Mattioli, R., Slingenbergh, J. & De La Rocque, S. (2008) Land cover and tsetse fly distributions in sub-Saharan Africa. *Medical Veterinary Entomology,* 22, 264-373.

Devisser, M., Messina, J., Moore, N., Lusch, D. & Maitima, J. (2010) A dynamic species distribution model of Glossina subgenus Morsitans: The identification of tsetse reservoirs and refugia. *Ecosphere,* 1.

Dicko, A. H., Lancelot, R., Seck, M. T., Guerrini, L., Sall, B., Lof, M., Vreyseng, M. J. B., Lefrançois, T., Fonta, W. M., Peck, S. L. & Bouyer, J. (2014) Using species distribution models to optimize vector control in the framework of the tsetse eradication campaign in Senegal. *Proceedings of the American Academy of Sciences*.

Ducheyne, E., Mweempwa, C., De Pus, C., Vernieuwe, H., De Deken, R., Hendrickx, G. & Van Den Bossche, P. (2009) The impact of habitat fragmentation on tsetse abundance on the plateau of eastern Zambia. *Preventive Veterinary Medicine,* 91, 11-18.

Guerrini, L., Bord, J. P., Ducheyne, E. & Bouyer, J. (2008) Fragmentation Analysis for Prediction of Suitable Habitat for Vectors: Example of Riverine Tsetse Flies in Burkina Faso. *Journal of Medical Entomology,* 15, 1180-1186.

Hay, S. I., Packer, M. J. & Rogers, D. J. (1997) The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International Journal of Remote Sensing,* 18, 2899 – 2930.

Holmes, P. (2013) Tsetse-transmitted trypanosomes – Their biology, disease impact and control. *Journal of Invertebrate Pathology,* 112, S11–S14.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. & Ferreira, L. G. (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment,* 83, 195-213.

Kitron, U., Otieno, L., Hungerford, L., Odulaja, A. & Brigham, W. (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, 371-380.

Matawa, F., Murwira, A. & Schmidt, K. S. (2012) Explaining elephant (Loxodonta africana) and buffalo (Syncerus caffer) spatial distribution in the Zambezi Valley using maximum entropy modelling. *Ecological Modelling,* 242, 189-197.

Matawa, F., Murwira, K. & Shereni, W. (2013) Modelling the Distribution of Suitable Glossina Spp. Habitat in the North Western parts of Zimbabwe Using Remote Sensing and Climate Data. *Geoinformatics and Geostatistics: An Overview,* S1.

Munang'andu, M. H., Siamudaala, V., Munyeme, M. & Shimumbo Nalubamba, K. (2012) A Review of Ecological Factors Associated with the Epidemiology of Wildlife Trypanosomiasis in the Luangwa and Zambezi Valley Ecosystems of Zambia. *Interdisciplinary Perspectives on Infectious Diseases,* 2012, doi:10.1155/2012/372523.

Odulaja, A. & Mohamed-Ahmed, M. M. (2001) Modelling the trappability of tsetse, Glossina fuscipes fuscipes, in relation to distance from their natural habitats. *Ecological Modelling,* 143, 183-189.

Parolo, G., Rossi, G. & Ferrarini, A. (2008) Toward improved species niche modelling: Arnica montana in the Alps as a case study. *Journal of Applied Ecology,* 45, 1410-1418.

Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling,* 133**,** 225–245.

Pender, J., Mills, A. & Rosenburg, L. (1997) Impact of tsetse control on land use in the semi-arid zone of Zimbabwe. Phase 2: Analysis of land use change by remote sensing imagery. *NRI Bulletin.* Chatham, Natural Resources Institute.

Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling,* 190**,** 231–259.

Phillips, S. J. & Dudik, M. (2004) A maximum entropy approach to species distribution modeling. *21st International Conference on Machine Learning.* Banff, Canada.

Pittiglio, C., Skidmore, A. K., Van Gils, H. A. M. J. & Prins, H. H. T. (2012) Identifying transit corridors for elephant using a long time-series. *International Journal of Applied Earth Observation and Geoinformation,* 14**,** 61-72.

Robinson, T., Rogers, D. & Williams, B. (1997) 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**,** 235-245.

Rogers, D. J., Hay, S. I. & Packer, M. J. (1996) Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. *Annals of Tropical Medicine and Parasitology,* 90**,** 225-241.

Rogers, D. J., Hay, S. I. & Randolph, S. E. (2000) Satellites, space, time and the African trypanosomiases. *Advances in Parasitology.* Academic Press.

Shereni, W. (1990) Strategic and tactical developments in tsetse control in Zimbabwe (1981–1989). *International Journal of Tropical Insect Science,* 11**,** 399-409.

Sibanda, M. & Murwira, A. (2012a) Cotton fields drive elephant habitat fragmentation in the Mid Zambezi Valley, Zimbabwe. *International Journal of Applied Earth Observation and Geoinformation,* 19**,** 286-297.

Sibanda, M. & Murwira, A. (2012b) The use of multi-temporal MODIS images with ground data to distinguish cotton from maize and sorghum fields in smallholder agricultural landscapes of Southern Africa. *International Journal of Remote Sensing,* 33**,** 4841-4855.

Terblanche, J. S., Clusella-Trullas, S., Deere, J. A. & Chown, S. L. (2008) Thermal tolerance in a south-east African population of the tsetse fly Glossina pallidipes (Diptera, Glossinidae): Implications for forecasting climate change impacts. *Journal of Insect Physiology,* 54**,** 114-127.

Van Den Bossche, P., De La Rocque, S., Hendrickx, G. & Bouyer, J. (2010) A changing environment and the epidemiology of tsetse-transmitted livestock trypanosomiasis. *Trends in Parasitology,* 26**,** 236-243.

Welburn, S. C., Coleman, P. G., Maudlin, I., Fãvre, E. M., Odiit, M. & Eisler, M. C. (2006) Crisis, what crisis? Control of Rhodesian sleeping sickness. *Trends in Parasitology,* 22**,** 123-128.