Where should the Muguga cocktail be used?

The distributions of Theileria parva, its hosts and vectors in Africa

William Wint and Henry Kiara

Environmental Research Group Oxford (ERGO)
c/o Department of Zoology, South Parks Road, Oxford, OX1 3PS, UK

International Livestock Research Institute (ILRI)
Old Naivasha Road, near Uthiru, Nairobi, Kenya 00100

December 2017

Acknowledgements

The authors are grateful to the many colleagues who have provided valuable advice, data, and information: Professor Graeme Cumming for his generosity with his extensive tick datasets; Dr Patrick W. Wagute of the Directorate of Resource Surveys and Remote Sensing in Nairobi and Dr Joe Ogutu for information on buffalo distributions in Kenya; Dr Marius Gilberts and Dr Tim Robinson for assistance with the cattle data; Dr Phil Toye of ILRI, Dr S. Ndungu and colleagues in the Muguga Veterinary Research Institute, Nairobi, and Dr Harry Oyo and colleagues at the Directorate of Veterinary Services, Kabete for advice in the epidemiology of buffalo derived *T. parva* in Kenya; Ms Rosekellen Njiru for so efficiently organising WW in Nairobi.

Table of Contents

List of Tables and Figures

Abstract

The parasite *Theileria parva* is carried by hard body ticks *(Rhipicephalus appendiculatus, R.zambeziensis)* and causes East Coast Fever in cattle and Corridor Disease in Buffalo throughout much of East and Southern Africa. If cattle come into contact with Buffalo derived strains of the parasite, they can also catch Corridor disease. Existing Infection and Treatment Methods using the Muguga cocktail may not prevent cattle infection where the two hosts coexist. Available distribution data for cattle, buffalo, the two vectors and the disease were collated and augmented with either spatial modelling using Boosted Regression Trees and Random Forest or habitat suitability masking, or both, to provide kilometre resolution maps of each target species. The spatial overlaps between wildlife hosts, vectors and the disease were identified, and the number of cattle in these overlap zones calculated. Approximately 60 million cattle are estimated to be within the areas where both vectors and disease are potentially present. Of these, depending on the assumptions used to define the contact areas, 3 - 5 million cattle could come into contact with Buffalo. Most of these are in Kenya, Uganda, Tanzania and Zimbabwe, which are each estimated to have between 0.5 and 1.5 million cattle in potential contact with infected Buffalo. The implications of these figures in the context identifying the next steps for treatment implementation are discussed.

1 Introduction and Objectives

The infection and treatment method (ITM) of immunisation against East Coast fever (ECF) has more than 95% efficacy against cattle derived *T.parva* infection. The evidence of the performance of the vaccine against Buffalo-derived *T.parva* is mixed. In some situations vaccination failed to protect cattle against buffalo-derived *T.parva* infection. In other areas, the vaccine appears to work perfectly even when cattle and buffalo graze together.

It is not clear if the reason for this difference is due to buffalo-derived parasites breaking through the immunity engendered by the Muguga cocktail strains and where the vaccine works is because somehow the buffalo derived *T.parva* stocks are protected by the vaccines. It is now known that buffalo derived parasites are more polymorphic than cattle derived parasites. It might also be that in the areas where the vaccine works, the proportion of buffalo derived parasites compared to cattle derived parasites is so low that the vaccine is able to protect against the cattle parasites which happen to be in the majority.

These issues require investigation to get a better understanding so that the vaccine could be improved. In the meantime, in order to protect the reputation of the vaccine, it is safer not to vaccinate in areas where cattle are known to interact with buffalo. Unfortunately these areas are not clearly known. Besides the number cattle at risk from the buffalo-derived parasites is not known to decide if it is worthwhile to invest in protecting those cattle.

This study aims to map the areas considered to be at risk and therefore to be avoided for ITM with the Muguga cocktail in its present form. This essentially involves identifying the areas where the spatial distributions of cattle, buffalo and the disease overlap.

The disease is transmitted between hosts – both domestic and wild – by tick vectors, and so the distribution of the disease is largely determined by the presence of infected ticks – primarily of the species *Rhipicephalus appendiculatus* and *R. zambeziensis.*

Estimating the geographic overlap therefore involves mapping each of the main component distributions (buffalo, cattle, and vectors) and estimating the degree to which these overlap. This requires a number of discrete stages, namely to:

- Review all available distribution data as well as the relevant literature in relation to risk of cattle infection from buffalo derived T. parva, including the efficacy of the Muguga cocktail as it is currently understood.
- Acquire and collate all relevant spatial data layers in a common format for the hosts and vectors, as well as the disease itself.
- It is already known that these data are not complete in that there are gaps in the known distributions which preclude assessing the degree of overlap within the entire area affected. It is therefore necessary to produce distributions that have such gaps filled. A well established method to achieve this is to use spatial modelling techniques to estimate the spatial distribution of the factors for which continuous, geographic estimates are not available or need updating.

- Combine these modelled distribution maps in such a way as to estimate risk to cattle from buffalo derived T. parva.
- Provide all spatial data to ILRI in a format suitable for further analysis by ILRI staff if required (see Section 11)

2 Review of Available Data

2.1 Epidemiological Overview

East Coast Fever (ECF) in cattle in caused by the protozoan parasite *Theileria parva*, and is widespread throughout East and Southern Africa (ILRI, 2017) . The disease in spread from host to host by tick vectors – primarily *Rhipicephalus appendiculatus* and *R. zambeziensis.*

The disease affects all cattle breeds including Zebu in pastoralist and agro-pastoralist production systems and ranches and improved dairy herds in the more intensive production systems. The parasite infects the bovine lymphocytes, causing a profound lymphoproliferation which can leads to mortalities of up to 80% in the more susceptible exotic (dairy) breeds. Economic losses can be very substantial –and ECF is widely regarded as the most serious animal health constraint to increasing the productivity of cattle in eastern, southern and central Africa. The parasite also causes disease in buffalo (*Syncerus caffer*) – known as Corridor Disease which causes much less severe symptoms than does ECF in cattle. Transmission can be of ECF between cattle, or Corridor Disease from buffalo to Cattle.

Sitt et al (2015) provides a comprehensive over view of the epidemiology of *T. parva*, from which much of the following is drawn . Though caused by the same species, the Corridor Disease parasites are significantly more genetically heterogeneous than those causing ECF, perhaps only a subset of the buffalo-maintained *T. parva* populations have crossed into cattle populations since the relatively recent introduction of cattle into eastern Africa (Hanotte et al., 2002).

Efforts to combat the disease have been ongoing for many decades. Direct treatments using antibiotics such as oxytetracycline can be used, but these do not, on their own, prevent re-infection Prevention is clearly better than cure and preventative methods can include tick control through acaricide pour-ons and the like, but the presence (and possible amplification) of the vectors within untreated buffalo populations allows the disease to persist in the environment and this remains a threat to domestic livestock (Walker et al, 2014) in and around areas where buffalo are found.

A more effective prevention technique was developed several decades ago, based on the Infection and Treatment Method ITM). This technique entails infecting cattle with a tripartite mixture of parasites derived from cattle and buffalo *T. parva* strains (the Muguga Cocktail) , and simultaneously treating the deliberately infected animal with antibiotics. The end result leads to the development of long term immunity in treated animals. The treatment is highly effective in most circumstances and extensive treatment campaigns based on the use of the Muguga Cocktail have been implemented in the region.

Though the technology has been available since the seventies, and the first vaccine productions started in the early nineties , implementation of these vaccination campaigns have, however, been limited by a series of essentially logistic constraints related to vaccine production and delivery (Kiara *et al*, 2016; Perry, 2016). These constraints, and the ways to overcome them, are now much better understood, and the practicality of substantially extending the use of the Muguga Cocktail is now much greater than has been the case to date.

There remain, however some technical issues which may limit deployment of the vaccine. Perhaps most important of these is the fact that the Muguga Cocktail may not protect against the disease carried by buffalo, probably because their relative heterogeneity means that the comparatively simple vaccine cocktail (though it does contain some buffalo derived elements) does not stimulate immunity to the entire range of buffalo parasites (Sitt, *op cit*).

The evidence to support this contention is equivocal: Homewood *et al* (2006) found that the vaccine was equally effective in areas with buffalo as in buffalo free locations, whilst recent studies have clearly demonstrated that vaccinated cattle can catch buffalo derived *T. parva* in some areas but not others close by (Sitt *et. Al., op. cit.,* Kiara, *in prep*). Whether this discrepancy reflects differing abilities of infected ticks to carry buffalo derived strains of the parasites – so that in some areas they cannot transmit buffalo derived strains to cattle – or variation in the susceptibility of cattle to the buffalo derived strains, or indeed regional variation in the buffalo strains themselves, remains unknown.

Until the answers to these uncertainties are found and the drivers of variation in cattle susceptibility to buffalo strains (whether geographic, immunological, or entomological) are sufficiently well understood to predict with a high degree of reliability it must be assumed that cattle can catch buffalo derived strains. Given the complexity of the system, the solutions are unlikely to be found in the short term, and in the mean time a strategy must be developed that allows the vaccine to be used in those areas where it works, and not in those where it may not. This approach is necessary primarily because a failure in vaccination will undermine stakeholders faith in the campaigns, and thus endanger the rate of uptake.' In essence, the precautionary principle therefore dictates that all areas which may have the disease and where cattle and buffalo are likely to come into contact should be excluded from vaccination campaigns.

3 Methodology of Estimating Required Distributions

Identifying these areas of overlap requires the combination of detailed estimates of cattle, buffalo, vector and disease distributions. In order for any such spatial combination to be effective, these distributions need to be compatible, *i.e*. cover the same area (extent); be available at the same level of detail (resolution) and refer to the same time periods.

The extent defined for this work is dictated primarily by the presence of the disease rather than the hosts which are found well beyond the region where the disease occurs. The tick vectors are also found beyond the disease range, but not to any great extent (See Section 4.2), and inclusion of their entire ranges within the study extent allows all the available data to be used to estimate

their distributions, which is likely to improve their reliability. The rather large extent also provide some flexibility when extract subset geographies. The defined extent is shown in Figure 1

Figure 1: Study Extent

The resolution used has been set at 1km, which is sufficiently detailed to be useful at a local scale, whilst being compatible with existing standard livestock distribution datasets (See 4.3.1). The time period is somewhat more complicated to define, given that some distributions may be determined by season and the available data relate to a broad data range. Pragmatism dictates that as much of the available data as possible are used to generate the necessary component distributions – more data provide better spatial accuracy of the outputs - which means that the outputs are in reality synoptic – *i.e.* are representative of a period of some years rather than any specific date. This also has a distinct advantage that synoptic data are less subject to anomalies caused by extremes or year to year perturbation and so are a more reliable prediction of a 'normal' situation.

The basic methodology for all these estimates is the same: assemble what data are available and if there are gaps then use established spatial modelling techniques produce complete predicted distributions at the required resolution. These techniques rely on establishing a statistical relationship between the presence or abundance of host, vector, or disease at known locations (the training data) and the values of a series of covariates at those same locations. These relationships are then applied to the covariate datasets, which are available for the entire area of interest, and can thus be used to generate a complete map with none of the gaps present in the observed distributions.

There are many statistical methods than can be used to implement these types of spatial analyses, and they can be use on either abundance or simple presence and absence data. With the exception of the livestock numbers (see Section 4.3.1) reliable abundance data are simply not available for large enough areas any of the host vector or disease distributions and necessity

therefore dictates the models must be for presence or absence. Two of the most widely used are Random Forest (RF) and Boosted Regression Trees (BRT) , both of which involve a degree of machine learning in the calculation of the predicted distributions. This study uses both these techniques, to produced separate presence models which are then combined into an ensembled model depicting probability of presence. The techniques are implemented using the VECMAP software suite which is also able to implement Zoned RF – producing predictions for defined analysis zones and integrating then into a single output. This tends to produce more reliable predictions as the relationships are tailored for each Zone. In this case the Zones used were the Livestock Production Zoned described in Section 4.3.1 below.

Presence absence modelling requires both presence and absence records in rough equal proportions. By their very nature, most survey efforts are designed to record presence rather than absence, and absence data are often rare and must be generated in one way or another. There are several ways to do this – typically either defining 'pseudo absences' which essentially rely on statistical ways of randomly generating absence points in localities where there are no presence records. The absence locations are determined only by their position relative to the known locations of presence rather than any characteristics of the locations themselves)

Alternative ways to define absences are to either identify areas that are known to be unsuitable based on the conditions that are thought to limit a species distribution (too hot, too wet, too bare of vegetation etc) or to assess the characteristics such as landuse of the locations with known presences and assign absence to land use types that have very few nor no presence records. These are the approaches used in this study and are detailed in each section below

These statistical distribution modelling techniques rely of the use of appropriate predictor covariates which typically include a wide range of environmental, climatic, demographic and agricultural parameters that are likely to determine the distributions of hosts or vectors or diseases. A comprehensive covariate suite was prepared which include an extensive set of remotely sensed imagery describing temperature, rainfall, and amount of vegetation density. A vast amount of satellite image data are now available, dating back to the mid eighties. This provides the choice of using climatic and environmental data from individual weeks, or years, and selecting either means or average or maxima. It is therefore necessary to reduce this archive to some manageable dataset that is nevertheless biologically meaningful . One such data reduction method is Temporal Fourier Analysis (TFA) that produces a series of climate indicators from an extensive time series of imagery, representing average levels, maxima, minima and seasonality. Such TFA covariates have been used for several years in spatial modelling (Scharlemann *et al,* 2008). This study uses a similar archive based on a 15 year 2000- 2015 time series of MODIS imagery (see Table 1), supplemented by elevation, human population and land use proportions. Details of files and formats are given in Section 11.3

Most of these categories comprise many different variables, so there are something in the order of 100 covariates available for each model. Using them all is likely to increase the chance of 'overfitting' *i.e.* the model then precisely replicates the input data. To reduce this false accuracy, each model was initially offered all the predictors available and then only the top ten were then used to provide the final outputs.

Table 1: Covariate Types Used in Spatial Modelling

Title	Description
MODIS DLST	Day-time Land Surface Temperature, 2000-2015 Fourier transformed dataset
MODIS NLST	Night-time Land Surface Temperature, 2000-2015 Fourier transformed dataset
MODIS NDVI	Normalised Difference Vegetation Index 2000-2015 Fourier transformed dataset
MODIS EVI	Enhanced Vegetation Index, 2000-2015 Fourier transformed dataset
MODIS CH3	Channel 3 Temperature, 2000-2015 Fourier transformed dataset
EarthEnv	Percentage Consensus Land Use categories.
GlobCover	Percentage Land Cover Categories 2009
Elevation	GMTED Minimum Elevation
RH	Minimum Annual Relative Humidity, 2015
Population	Worlpop Human Population
TAMSAT	Monthly precipitation, 2016

Each of the following Sections contain descriptions of the available data and any procedures used for assigning absences which are then followed by the model outputs.

4 Distributions

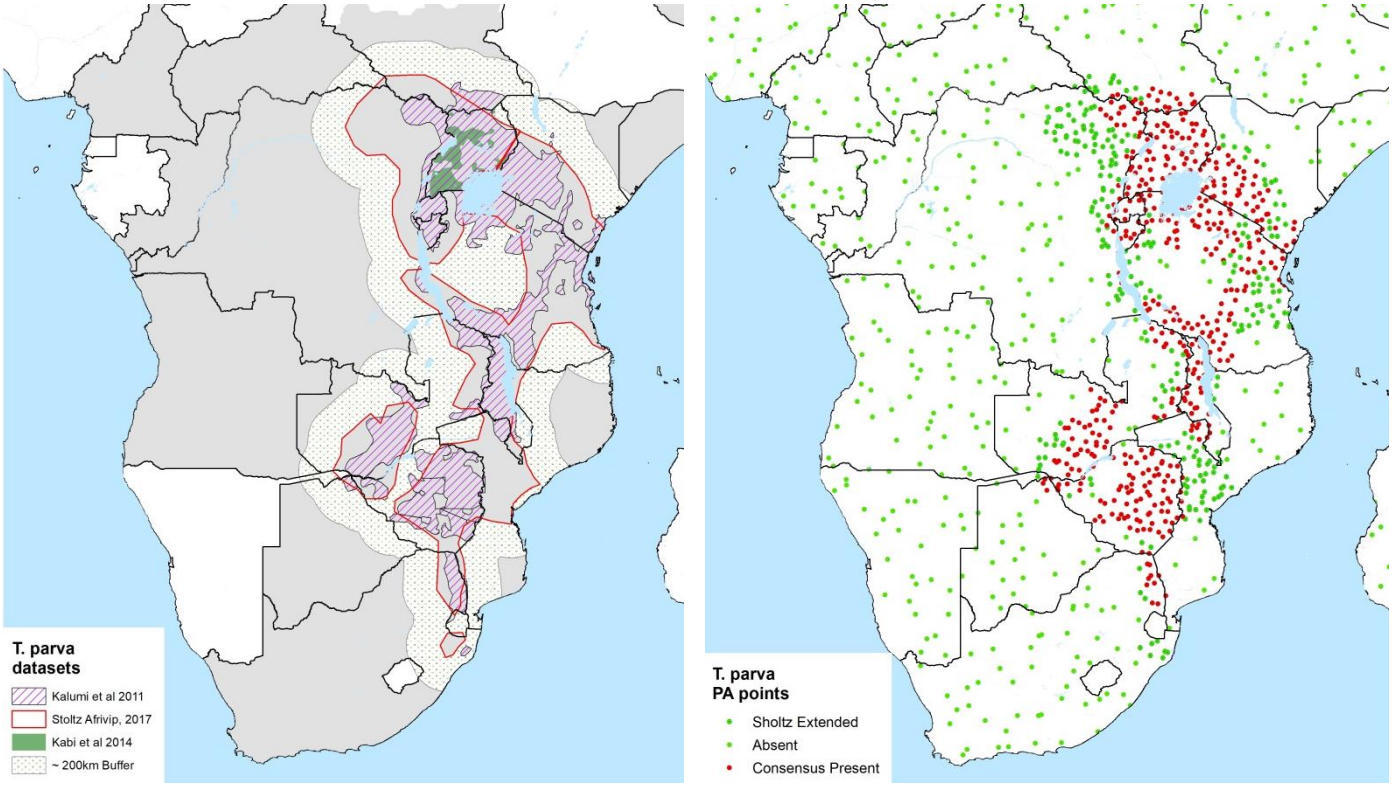
4.1 The Disease

Norval *et al* (1992), listed 16 countries for which ECF had been recorded, to which should be added Angola De Garcia and Serrano (1971), Cameroon and the Democratic Republic of the Congo (DRC) (Kiara, *pers. comm.*). The complete list is therefore 18 countries: Angola, Botswana, Burundi, Cameroon, Central African Republic, Congo, Democratic Republic of the Congo, Kenya, Malawi, Mozambique, Rwanda, Somalia, South Africa, South Sudan, Swaziland, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe.

Whilst there are a number of surveys for Corridor Disease and East Coast Fever, they largely relate to very small areas, and so are at too fine a scale to be used effectively at a sub-continental level. There are two relatively recent publications that collate all available reported information into a full coverage based on what amounts to informed expert opinion, namely Kalumi, Losson & Saegerman (2011) and Stoltz (2017) derived from Lawrence et al (2005a, b). Whilst comparatively similar in general terms, the two coverages differ in some important details, for example an extension of the distribution into eastern DRC and southeast Tanzania as shown in Figure 2.

Though disease distributions have not been systematically surveyed and mapped for the entire study area, there are a few individual studies which provide information about particular countries (Kabi *et al* , 2014 for Uganda; Proceedings of an ILRAD workshop, 1989 for several countries).

Figure 2: Distribution of Theileriosis a) published maps (left); b) presence and absence points assigned (right).



Countries where the disease has been recorded are shaded grey

As there are no large scale prevalence or incidence data available from these datasets, the only alternative is to model presence and absence. There are however no specific absence locations in the training data, nor are there presence points for which environmental conditions can be extracted and evaluated to infer absence where such conditions are different. This leaves the single option of generating some form of 'pseudo absence' based on location relative to the regions of known presence.

Given the discrepancy between the training data sets, the main body of absence points were generated in the regions beyond a buffer zone that bounded the combined extent of the training data. Absences were also generated within the buffer but at a lower spatial density, thereby creating a gradient in assumed likelihood of absence as distance from known presence increases. No absences were generated in locations where either training data set showed positive.

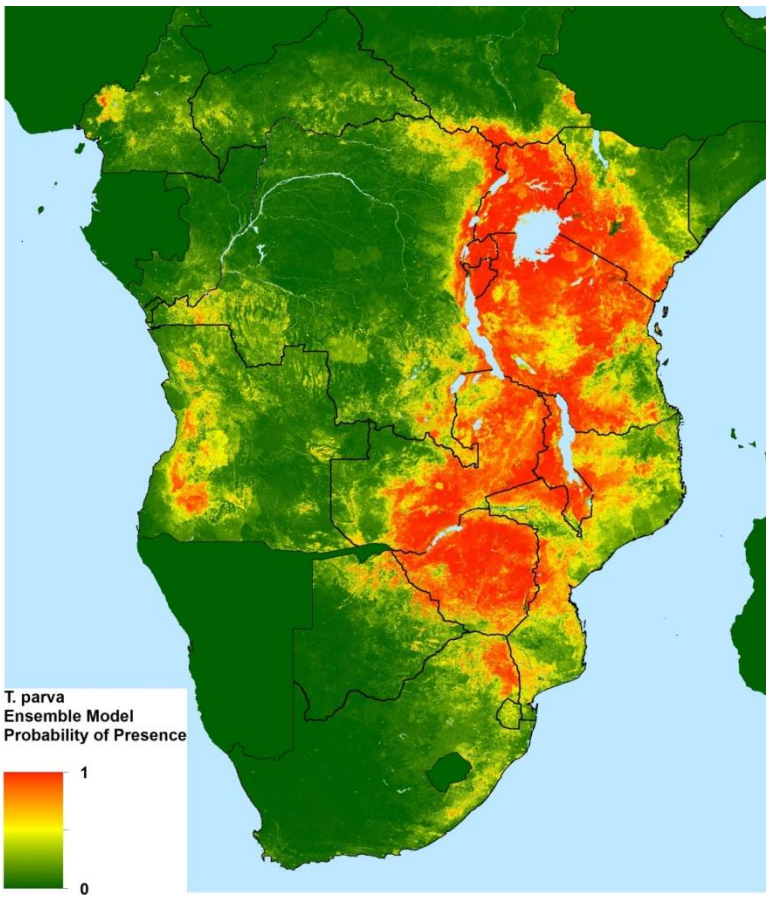
Two models were initially run, each with the same absence points (as implied by both training data sets) but with presences derived from only one of the training datasets. These effectively represented minimum and maximum 'known' distributions – named Consensus and Extended distributions. The main predictors of the BRT modelled distributions for these two models (shown in Table 2) are Rainfall in March and November, minimum relative humidity the seasonality and range of temperature related parameters, elevation, and human population level. None of the vegetation related factors were represented in the top ten predictors.

Table 2: Predictors, *T. parva* BRT models

Importance Max Model	Variable	Importance Min Extent	Variable
14.79	ECTSAT0316 March Rainfall	12.52	ECTSAT0316 March Rainfall
11.32	EC011503A2, Ch 3 temp, Amp2	11.68	EC011503A2: Ch 3 temp, Amp2
9.06	ECTSAT1116, Nov Rainfall	10.72	ECMN30GRD Elevation
5.06	ECWPPDN15A Human Popn Density	10.64	ECTSAT1116 Nov Rainfall
4.72	ECTSAT0416 April Rainfall	7.04	ECWPPDN15A Human Popn Density
4.71	ECRHMIN15 MInimum RH	4.28	EC011503P3 Ch3 Phsae 3
3.43	ECMN30GRD Elevation	4.14	EC011507P3 Day LST Phase 3
3.08	EC011503P3 Ch3 Phase 3	4.13	ECRHMIN15 Minimum RH
2.91	EC011515P1 EVI Phase 1	2.90	EC011508A1 Bight LST Amp1
2.57	EC011507P1 Day LST Phase 1	2.29	EC011507A2 Day LST Amp 2

The predicted distributions are shown in Additional Figure 11, (Section 8), and both reflect the input datasets well. Both (especially the extended distribution model) also predict the disease to be fairly extensive in southwest Ethiopia, from where it has not been reported. Following discussion with the relevant experts at ILRI and Kenyan Veterinary Agencies, and given the fact that Land Cover, Land Use and Environmental conditions suitable for the tick and for cattle only extend for a very limited distance into DRC, it was decided to rely on the consensus distributions as the more realistic of the two alternatives, and so an additional RF model was run on this dataset. The ensemble output is shown in Figure 3. Note that has been masked in this figure to remove predicted presences in countries where the parasite has never been reported.

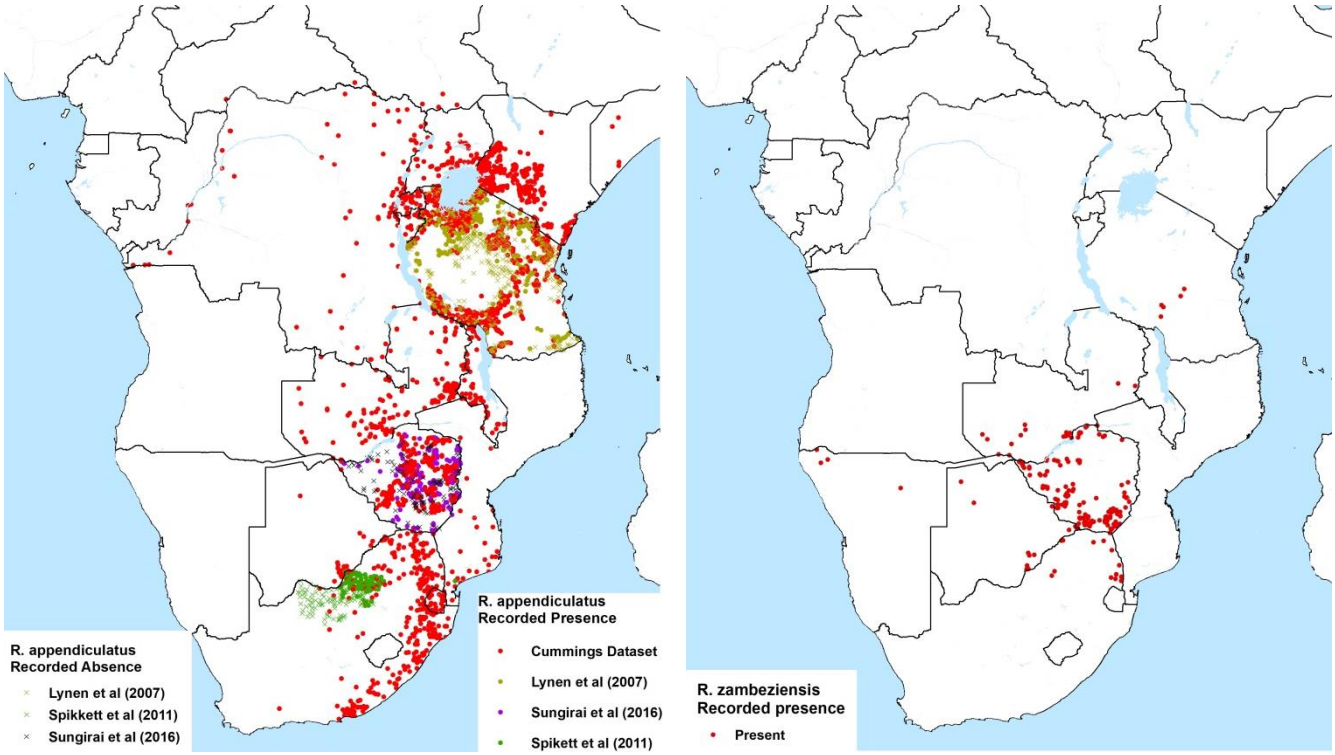
Figure 3: Predicted Parasite Distribution



4.2 The Vectors

There are two primary vectors of *T. parva* – both hard ticks: *Rhipicephalus appendiculatus* and *R zambeziensis.* The most extensive datasets for the distribution of these vectors was compiled by Cumming (1999), which the author kindly made available to this study and as the most comprehensive data currently available form the backbone of the vector distributions used in this work. There have however be quite a few more recent surveys of the vector distributions which have been used to update and enhance the Cummings data. For *R. appendiculatus*, these include Gachohi *et al* for Kenya. Lynen *et al* (2007) for Tanzania, Sungarai *et al* (2016) for Zimbabwe, and Spikett *et al* (2011). For *R. zambeziensis,* Cumming's data also formed the backbone of the available information. The recorded presence of both species are shown in Figure 4

Figure 4: Recorded Vector Distributions: a) *R.appendiculatus*; b) *R. zambeziensis*



These distributions are dominated by presence records so, as set out in the methods, they require absence records to be generated to enable presence absence modelling by the chosen techniques. These where generated for both species in two ways. Firstly it was assumed that the species were absent if more than 220 km (2 degrees) from a known or assumed presence. In this case the presence absence boundaries were taken from an additional set of vector distribution maps was published online by Madder, Horak, and Stoltz on the African Veterinary Information Portal. These appear to be modified and buffered versions of the Cumming distributions and as such depict a broader distribution. As such the absences derived from them are conservative.

The second way used to generate absences used the used the ecoclimatic conditions at each of the recorded presence points and identified those ecoclimatic categories where presence points were absent or very nearly so. The figures relative frequencies were calculated as percentages which were compared to the proportion of each ecoclimatic category within the bounding coordinates of the presence records. Thus, if only 1% of the presence points were in land use

category 5, but that category covered 25% of the presence region, then that LU category was defined as unsuitable for the vector. Conversely if LU category 10 covered 25% of the presence area but 10% of the presence records were in that LU type, then that LU class could not be defined as unsuitable, and so could not be defined an Land use category where the vector could be assumed to be absent. A category was defined as unsuitable if ratio between the two percentages was less than 0.25. The parameters used in this way are shown in Table 3 along with the categories identified as unsuitable.

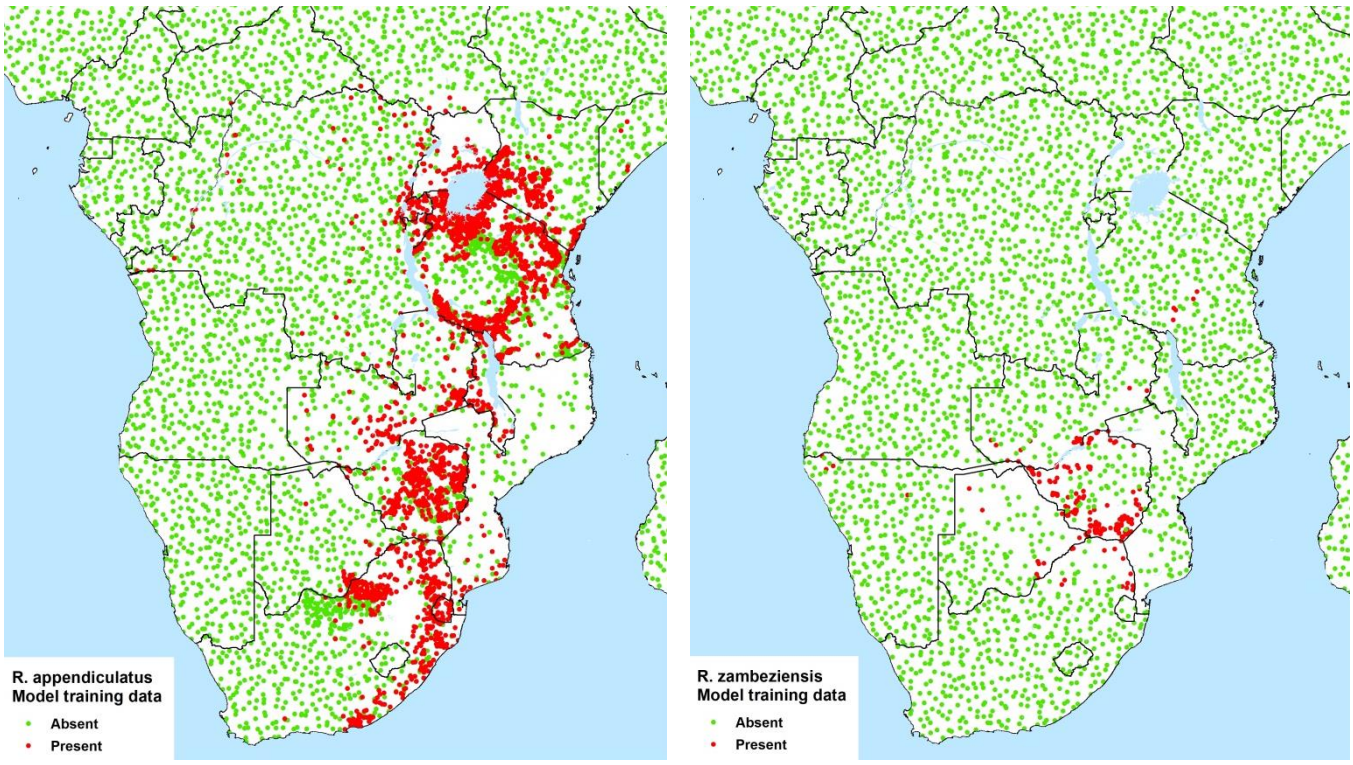
Table 3: Parameters used to define unsuitable conditions

Parameter	**Unsuitable conditions defined**	
	R. appendiculatus	***R. zambeziensis***
Elevation	>2500m	>1100
Maximum Normalised Difference Vegetation Index	< 0.45	< 0.35
Average daytime temperature	<26C	<28C
Total annual Rainfall	>1100mm or < 350mm	>850mm < 250mm
Globcover See below for classes	40,60,70,100,170,180,200 – 230	11,14,20,40,50,70,90,100, 120,170, 200, 220 230

Globcover class codes are as follows: 11=Post-flooding or irrigated croplands (or aquatic); 14=Rainfed croplands. 20=Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%); 30=Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%) 40=Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m).50=Closed (>40%) broadleaved deciduous forest (>5m).60=Open (15-40%) broadleaved deciduous forest/woodland (>5m).70=Closed (>40%) needleleaved evergreen forest (>5m).90=Open (15-40%) needleleaved deciduous or evergreen forest (>5m).100=Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m).110=Mosaic forest or shrubland (50-70%) / grassland (20-50%).120=Mosaic grassland (50-70%) / forest or shrubland (20-50%) .130=Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m).140=Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses).150=Sparse (<15%) vegetation.160=Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily) - Fresh or brackish water.170=Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water.180=Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water.190=Artificial surfaces and associated areas (Urban areas >50%).200=Bare areas.210=Water bodies.220=Permanent snow and ice.230=No data (burnt areas, clouds,…).

The resulting presence & absence datasets used to train the spatial models are shown in Figure 5.

Figure 5: Presence and Absence points defined for: a) *R.appendiculatus*; b) *R. zambeziensis*



The ensemble model outputs for each species are shown in Figure 6. They both reflect the training data well, with specificities and sensitivities in excess of 93% and 85% respectively. As is often the case with such models, there are a number of noticeable false positive areas – southwest Angola and southwest Ethiopia for *R. appendiculatus,* and along the Caprivi strip and north-west Mozambique for *R. zambeziensis.* Whilst these might indeed be prediction errors, they do represent suitable niches for the vectors and surveillance in these areas might reveal hitherto unknown populations. The predicted presence for *either* vector species,(i.e. where the probability of presence is greater than 50% for either) is shown in Additional Figure 12

Figure 6: Probability of Presence Ensemble Models: a) *R.appendiculatus*; b) *R. zambeziensis*

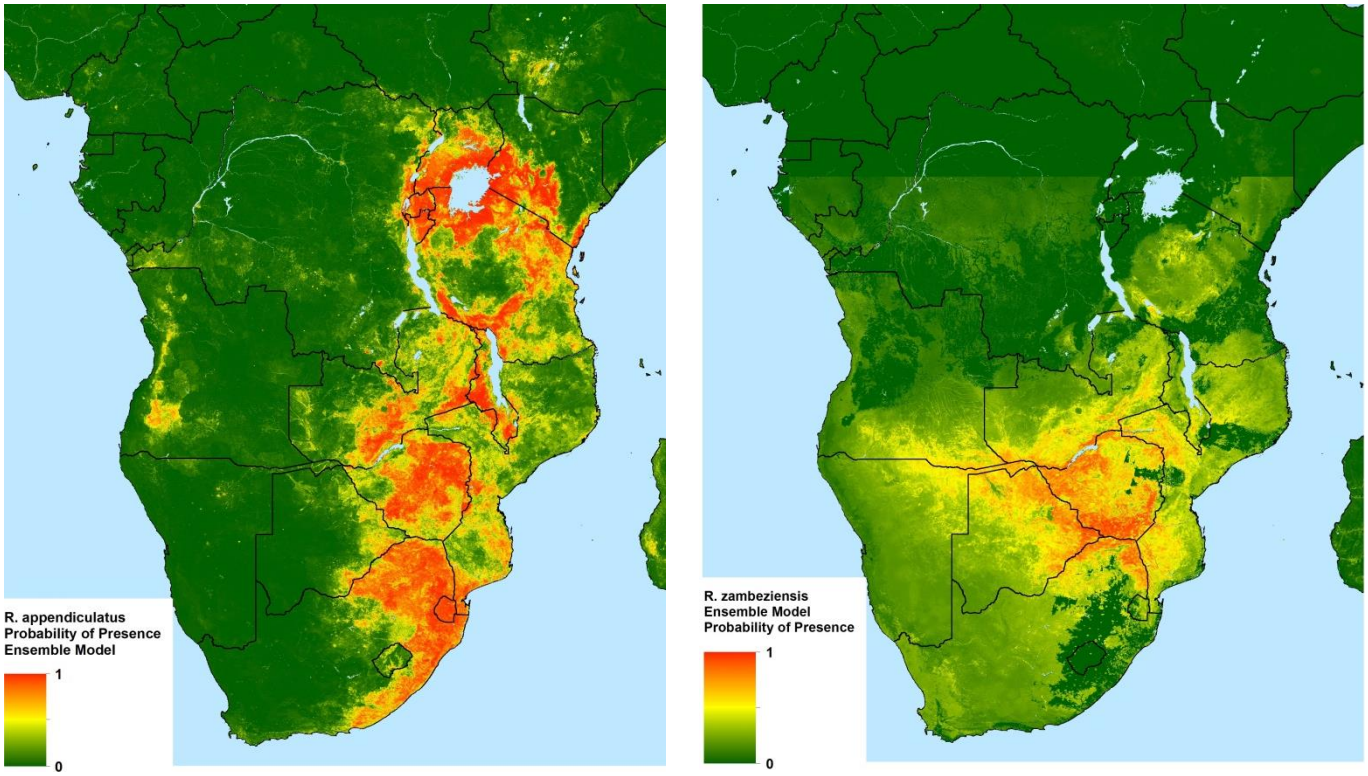


Table 4: Top 10 Predictors, Vector Models

R. appendiculatus		***R. zambeziensis***	
Importance	**Variable**	**Importance**	**Variable**
25.18	Human population density, 2010	17.67	*Enhanced Veg Index, Phase 2*
20.37	Rainfall January 2016	17.26	Elevation
10.65	*Enhanced Veg Index, Phase 2*	14.73	*Normalised Diff Veg Index, Phase 2*
9.15	*Daytime Land Surface Temp, Phase1*	10.74	*Normalised Diff Veg Index, Phase 1*
7.67	Rainfall March 2016	9.80	*Night-time Land temp, Amplitude 1*
6.73	Rainfall November 2016	8.94	Distance to Water
6.16	Minimum Relative Humidity, 2015	7.10	*Minimum Day-time Land Temp*
5.98	*Mean Night-time Land Temp*	6.74	*Night-time Land temp, Phase 1*
4.42	*Max Night-time Land Temp*	4.54	*Night-time Land temp, Phase 3*
3.63	Rainfall August2016	2.43	*Enhanced Veg Index, Phase 2*

Variable in italics are from the Temporal Fourier Analysed MODIS datasets (See Table 1)

The top ten model predictors for each species are quite contrasting, and are shown in Table 4. These are human population, several rainfall parameters and Relative Humidity, night time temperature metric and seasonality of Enhanced Vegetation Index for *R. appendiculatus*; and a predominance of vegetation seasonality metrics, night time temperature levels, and elevation for *R. zambeziensis*

4.3 The Hosts

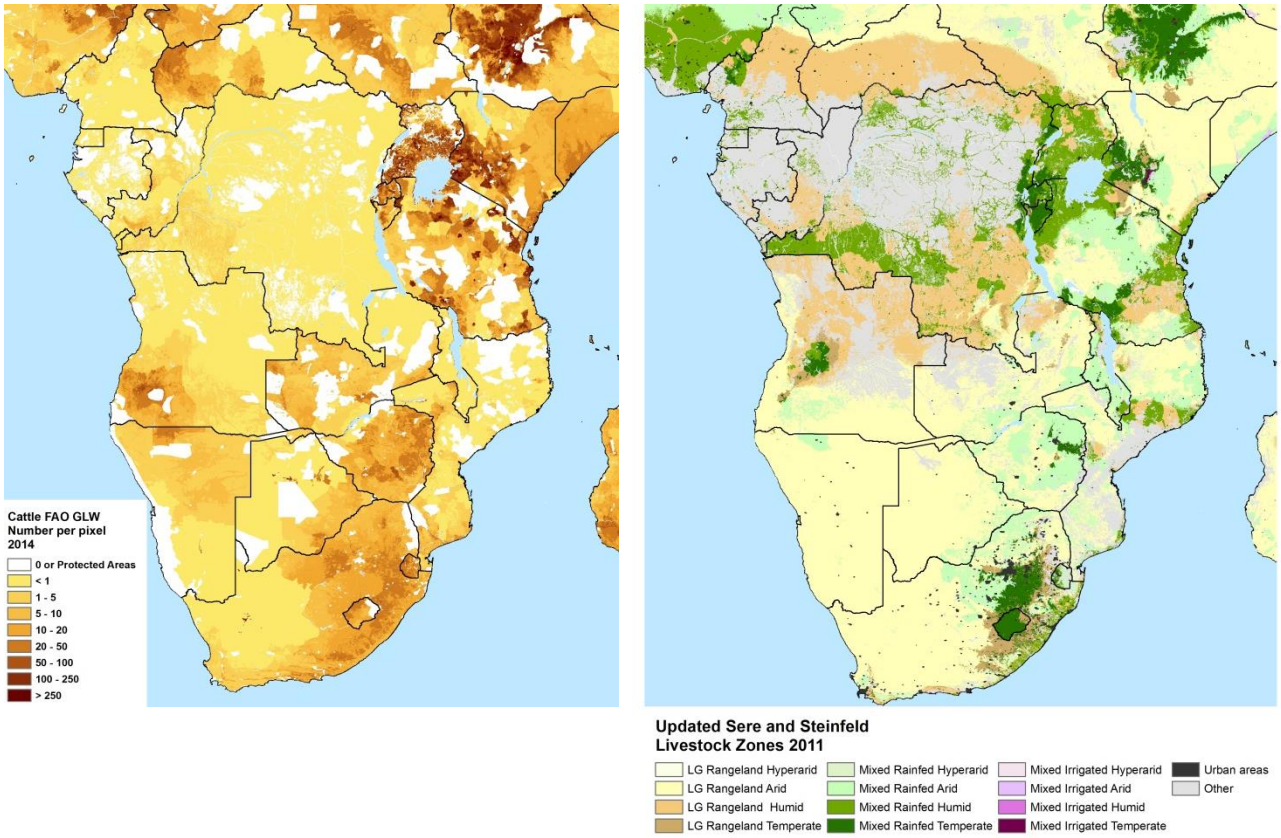
Two host species are of concern for this study: Cattle and Buffalo

4.3.1 Cattle

Cattle distributions are available from the Gridded Livestock of the World dataset (FAO, 2017). First produced in 2007 (Wint and Robinson, 2007), this global standard dataset has now evolved to version 2, and provides cattle distributions for 2010 at a resolution of 1km. These products are, in fact, a synoptic average of the dates of the reported data sets from which the outputs are derived. The date assigned to the version reflect the dates of the FAO national population estimates that are used to standardise the values so that national totals match official numbers. This structure makes it possible to 're-calibrate' the 2010 dataset to match national totals from different years. For this study, the population figures – available as both densities per square kilometre and numbers per image pixel (0.00833 degrees) were converted to 2014 values – the latest dates held within the FAO archives. Figure 7a shows cattle distribution to be widespread throughout the likely distribution of ECF.

Cattle production systems are classically divided into categories such as pastoral, agro-pastoral, mixed, Intensive. Whilst these classes have been estimated and mapped on the basis of ecology and economic contribution by livestock for the Horn of Africa (Cecchi *et al* 2010) they are not available for the rest of the continent. The only livestock production classification of the entire study area is by Sere and Steinfeld (1996), updated by Robinson *et al* (2011) shown in Figure 7b

Figure 7: a) Cattle Numbers per Pixel, 2014; b) Livestock Production Zones



4.3.2 Buffalo

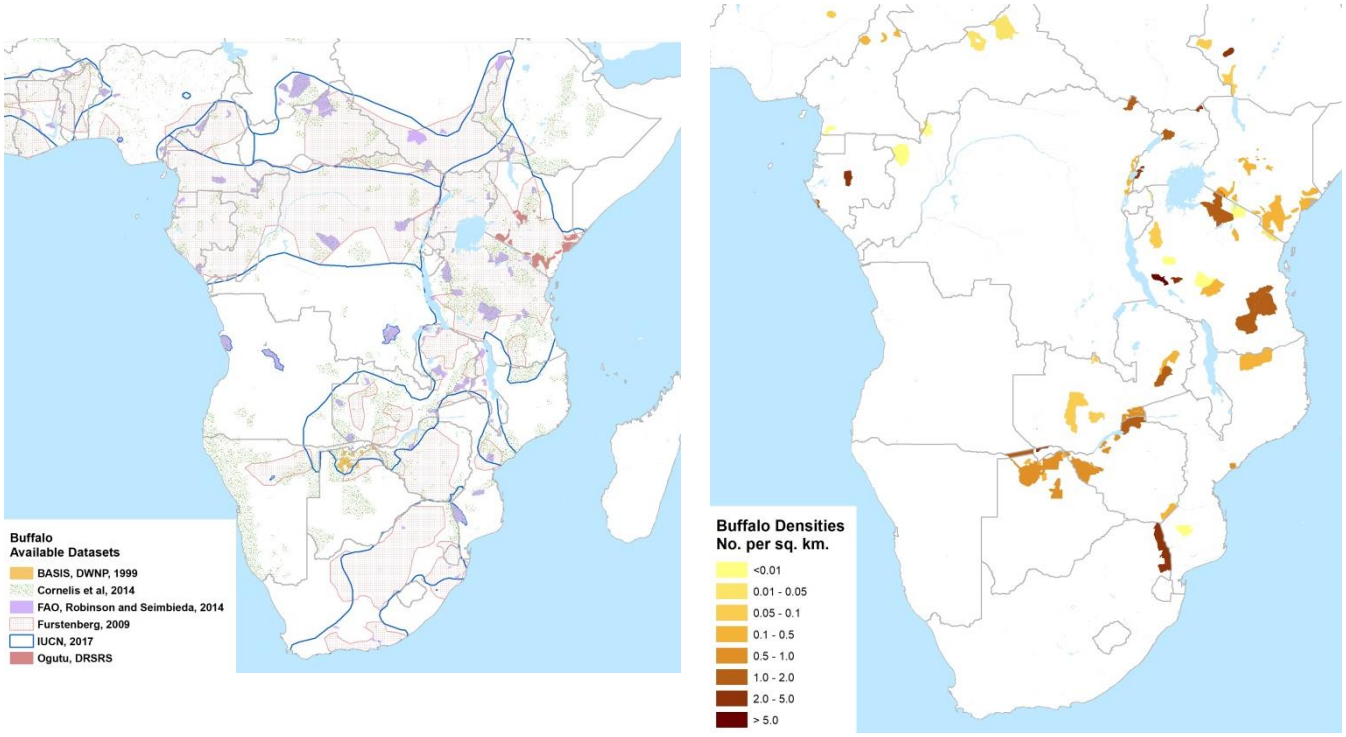
Despite being one of the 'big five', there appear to be remarkably few high resolution distribution data for Buffalo as a whole or for any of its three races available in the public domain. This may be partly because Buffalo numbers are changing rapidly which makes it difficult to keep track of population size. There are, however, a number of references that provide buffalo numbers by protected area or country. The seminal publication by East (1998) gives a Cape buffalo population of 548,00, of which 142,000 (26%) were found outside protected areas – mostly in Tanzania, Zimbabwe, Botswana, Zambia and Kenya. This has been followed by a most comprehensive by Cornélis *et al*, (2014) which is encyclopaedic in the detail it provides on both habits and recorded numbers within protected areas. This publication give a Cape buffalo population of 473, 000. A second comprehensive data set of buffalo numbers in National Parks was produced by FAO (Robinson and Siembieda, 2011) which has been made available to this study.

There are two other authoritative continental datasets for this host: the range boundaries produced and kept reasonably current as part of the Mammal Database (IUCN, 2017) hosted by the International Union for Conservation and Nature (IUCN). A map was also produced by Furstenberg, (2009), but was only published online and is not substantiated by data sources or any explanations and so, whilst potentially accurate, is shown in Figure 8a below for information only.

There are in addition a number of local or regional estimates of wild buffalo – notably the aerial surveys conducted by the Department of Wildlife and National Parks of Botswana; the Kenyan Directorate of Resource Surveys and National Parks, and the Tanzanian Wildlife Research Institute. Population numbers from these surveys were converted to mean densities with the survey

boundaries in order to be compatible with the Cornélis datasets, so that the processing and analyses could be standardised, as illustrated in Figure 8b.

Figure 8: Buffalo Distributions: a) Available Datasets; b) Densities Used



Buffalo are also farmed widely, especially in South Africa, in many if not most of the hundreds of Private Game Reserves and ranches in the country. The locations of these are not in the public domain and are likely to require many months of painstaking detective work to unearth Fortunately, a study by Mbizeni *et al* (2013) strongly suggests that the only threat to cattle is from buffalo in farms near the Kwa-Zulu Natal Reserve, within an ECF controlled area, in the far North East of the country, and the rest of the farms do not represent a threat. They have, not, therefore been included within this study.

The combination of these density data provide a good basis for estimating numbers within protected areas, and in some countries (such as Botswana and Kenya) outside the parks. A comparison of the numbers at country level with those protected areas from the figures provided by Cornelis (*op. cit*) suggest that the numbers outside protected areas are in fact minimal, though this conclusion remains at odds with East's (1998) historical perspective.

Clearly not all the areas within Protected Areas can support buffalo, and it is therefore necessary to identify the suitable habitats within the park boundaries. There are numerous studies of Buffalo ecology and the determinants of its distribution, movement and behaviour. Distributions have been inferred from habitat suitability within the IUCN boundaries by the African Mammal Databank (IEA, 1998). This latter database has been updated in recent years, but is currently being withheld from the public domain.

The published literature (e.g. Prinz, 1996; Kingdon, 1997) agrees that there are a number of constraints to buffalo distributions most obvious of which is the animal's preference for staying

close to open water. Reports of the threshold distance vary to some degree but the consensus appears to be that this species is usually not found beyond 20 km and very rarely beyond 25km, from open water. Descriptions of habitat preferences vary according the sub-species– Forest Buffalo prefer thicker vegetation, whist the Cape Buffalo is more frequently found in Open Wooded Savannah. Given the distribution of the *T. parva*, which is not found in the heavily forested areas, this study has included all Woodland categories as suitable habitats. Forest percentage was obtained from the 100m resolution CCI Land Cover dataset for 2015 (ESA, 2017).

The presence of water has been extracted from several sources: a) a 20m resolution Land Cover dataset (ESA, 2016) which are sufficiently detailed to provide a good indication of even relatively small sources of open water; b) from the small water body archive produced as part of the ESA Copernicus datasets (VITO, 2017). This is supplied as ten day imagery at 300m and so allows the calculation of the proportion of a year for which each pixel is covered in water; c) the Hydrosheds 15s resolution river courses dataset (USGS, 2006).

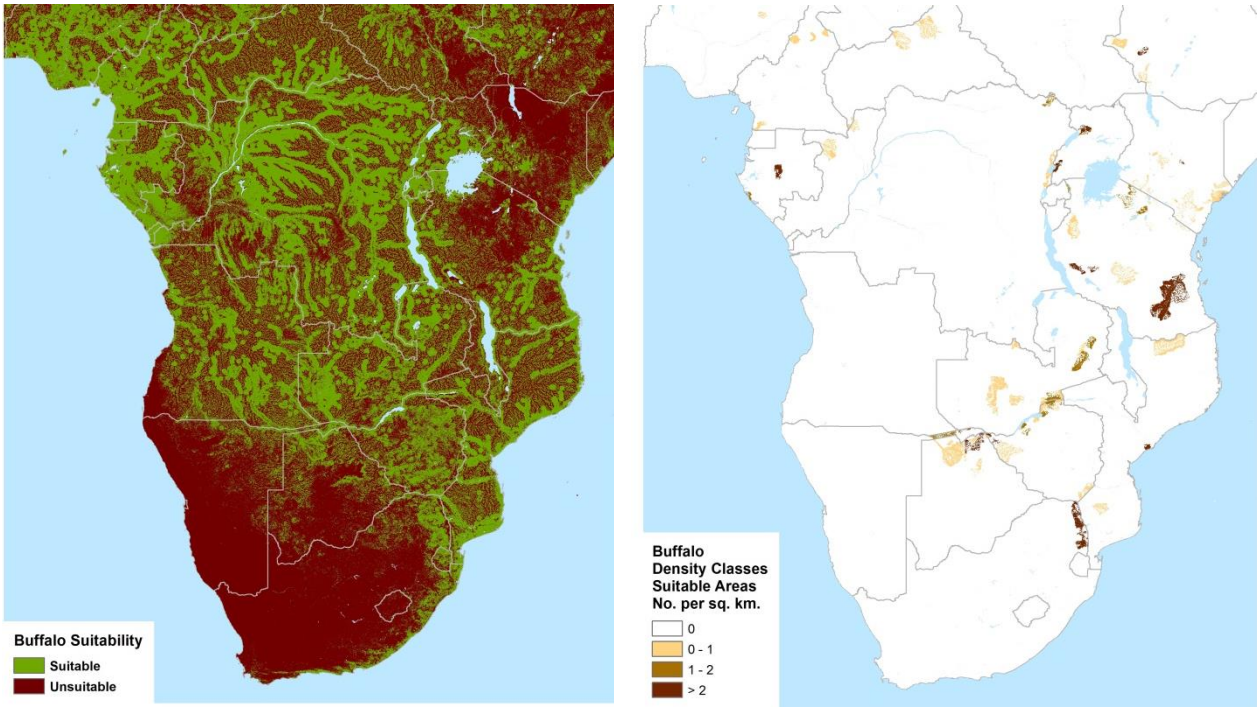
These datasets have been combined to define habitat suitable for buffalo as land with more than 10% forest and within 0.13 degrees (= approx. 15 km.) of river courses or 0.16 degrees of permanent water as detected by satellite imagery. This is illustrated in Figure 9a.

If populations do indeed persist outside protected areas, there is a case for attempting a spatial model to locate them using the standard approach, as for vectors and livestock, based on known data and inferred habitat suitability and/or unsuitability. The dependence on water is especially useful in this context and could be used to mask the species distributions.

Even though farms have been discounted (see above), there are some issues that make modelling more difficult, most especially the fact that the remaining distributions of buffalo populations outside parks are largely unknown, and are likely to occupy a very limited fraction of the suitable niche that a model would identify and there are no obvious ways to mask out areas from which the

After due consideration and expert consultation it was decided to adopt the following approach. Rather than producing a spatial model of buffalo numbers, the suitability map based on habitat and water availability has been used to mask out the areas within which density figures are available within the protected areas or surveyed areas. This distribution is shown in Figure 9b.

Figure 9: Buffalo: a) Suitable Habitat; b) Suitable Habitat Density Classes



5 Numbers of Cattle at risk.

The previous sections have described the production of the data layers needed to locate where cattle are at risk of transmission of buffalo derived *T. parva.* These are the distributions of the disease, cattle; the two disease vectors, and the Cape buffalo together. These have been combined to identify the overlapping regions.

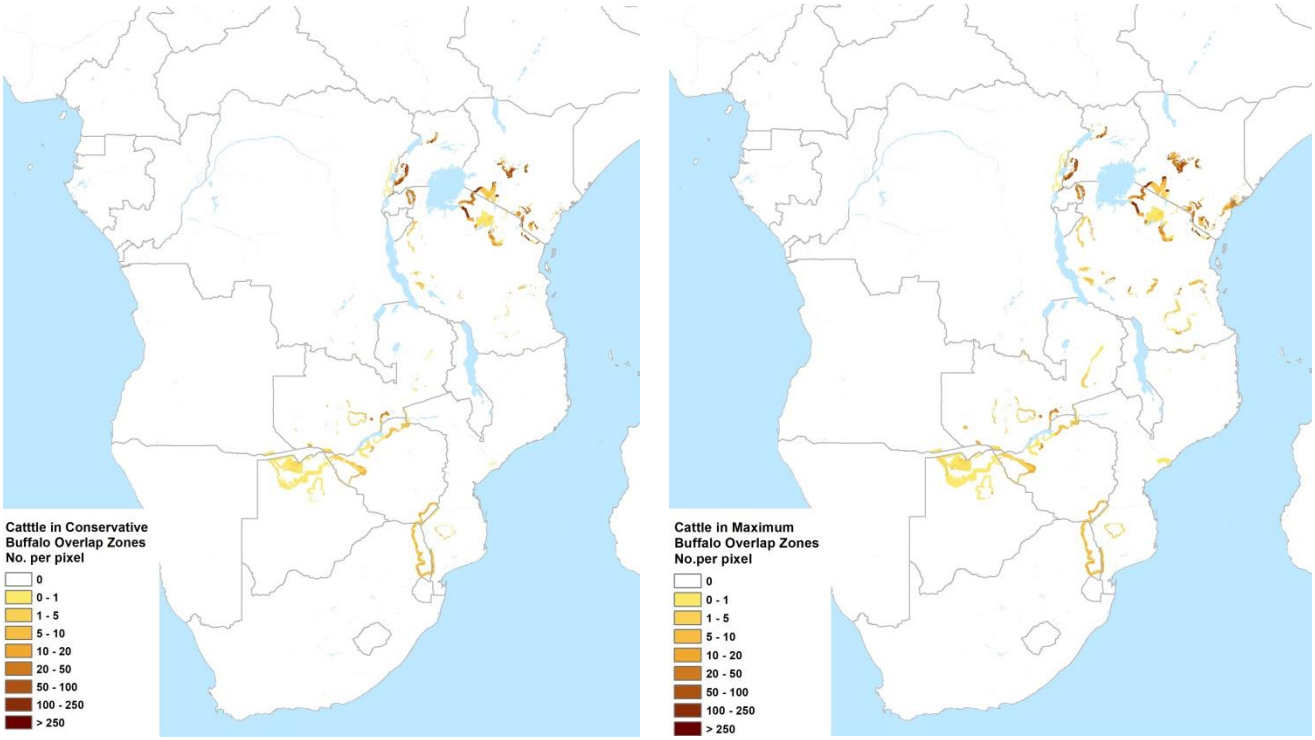
The overlaps between disease and the vectors, calculated using a 50% probability of presence as the threshold for Presence are shown in Additional Figure 13a, together with the cattle in these areas in Additional Figure 13a.

Discussion with ILRI staff and professional colleagues in the Directorate of Veterinary Services and the Veterinary Research Institute confirmed that a simple combination of the four components produced (disease, vectors, cattle, buffalo) was not sufficient and that a more sophisticated approach was required. In particular it was recognised that cattle moved substantially during their daily grazing routines, and that the livestock could therefore interact with ticks carrying buffalo derived disease some considerable distance from the herd's home range. As a result it was concluded that the 'risk' zone around buffalo distributions should include a buffer, the size of which depends on the numbers (densities) of buffalo present.

A buffer zone between cattle and buffalo was therefore calculated to take account of potential cattle movement to grazing areas beyond the edge of the herd. Cattle move up to some 30km a day – which was therefore defined as the maximum buffer distance which was adjusted according to the estimated buffalo density as follows: up to .5 per sq. km = 10km; 0.5 – 1 sq. km = 15km; 1 – 3 sq. km = 20km, above 3 sq. km = 30km. This provides an estimated 'buffalo derived risk zone' for which cattle populations can be estimated and is shown in Additional Figure 14

The main high–level metrics required from the study were identified as the numbers of cattle in a) the areas with the vectors and the disease estimated to be present; and b) within the 'buffalo derived risk zones' in areas with the parasite and the vectors present. The definition of these geographical overlaps using the modelled predictions of vectors and disease depends the probability thresholds used to define presence. Two were used: i) the standard probability of 0.5, representing the likely presence of each modelled component; and i) a reduced probability of 0.25 representing a less conservative (akin to a worst case) likelihood of presence. A similar approach can be taken to define significant cattle populations, and two thresholds were defined: under 5/sq. km., and under 1/sq.km., the latter therefore including very sparse livestock populations, likely to be existing in marginal conditions. Combining these two sets of thresholds has provided two overlap Zones: a) a "Conservative" Zone based on presence define as > 0.5 probability and cattle as < 5/sq.km; and b) a "Maximum" overlap defining presence as >0.25 probability and cattle as < 1/sq.km. The numbers of cattle per pixel in these Zones are shown in Figure 10.

Figure 10: Cattle in Buffalo derived risk zones a) Conservative; b) Maximum



The FAO Gridded Livestock of the World estimate there to be approximately 120 million cattle in the 18 countries ECF affected countries listed in Section 4.1. The estimated numbers of cattle in the various risk categories are shown in Table 5.

Table 5: Cattle numbers in different Overlap Zones

Country	Cattle No (m) 2014 (FAO)	Disease and either vector		Disease and *R. zambeziensis*		Disease and *R. appendiculatus*		Buffalo, Disease and either vectors			
								Conservative Thresholds		Marginal thresholds	
		Number (m)	*% total*	Number (m)	*% total*	Number (m)	*% total*	Number (m)	*% total*	Number (m)	*% total*
Angola	4.90	0.61	*12.5*	0.01	*0.3*	0.60	*12.2*	0.00	*0.0*	0.0	*0.0*
Botswana	1.80	0.12	*6.8*	0.12	*6.6*	0.05	*2.8*	0.02	*1.1*	0.0	*1.7*

Burundi	0.60	0.32	*52.3*	0.00	*0.0*	0.32	*52.3*	0.00	*0.0*	0.0	*0.0*
Cameroon	5.95	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.0	*0.0*
Central African Republic	4.35	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.0	*0.0*
Congo	0.34	0.00	*0.1*	0.00	*0.0*	0.00	*0.1*	0.00	*0.0*	0.0	*0.0*
Democratic Republic of the Congo	0.95	0.01	*1.3*	0.00	*0.0*	0.01	*1.3*	0.00	*0.2*	0.0	*0.4*
Kenya	17.81	12.11	*68.0*	0.00	*0.0*	12.11	*68.0*	1.02	*5.7*	1.4	*7.9*
Malawi	1.32	0.03	*2.3*	0.00	*0.1*	0.03	*2.3*	0.00	*0.0*	0.0	*0.0*
Mozambique	1.59	0.45	*28.6*	0.18	*11.2*	0.38	*23.8*	0.05	*3.2*	0.1	*3.7*
Rwanda	1.14	1.11	*97.3*	0.00	*0.0*	1.11	*97.3*	0.10	*8.5*	0.1	*9.3*
Somalia	4.90	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.00	*0.0*	0.0	*0.0*
South Africa	13.92	0.82	*5.9*	0.25	*1.8*	0.81	*5.8*	0.09	*0.7*	0.1	*0.7*
South Sudan	11.82	0.01	*0.1*	0.00	*0.0*	0.01	*0.1*	0.00	*0.0*	0.0	*0.0*
Swaziland	0.62	0.07	*10.8*	0.00	*0.4*	0.07	*10.8*	0.00	*0.0*	0.0	*0.0*
Uganda	13.62	9.37	*68.8*	0.00	*0.0*	9.37	*68.8*	0.66	*4.8*	0.7	*5.3*
United Republic of Tanzania	25.80	17.25	*66.9*	0.16	*0.6*	17.14	*66.4*	0.90	*3.5*	1.1	*4.4*
Zambia	4.09	1.04	*25.4*	0.31	*7.5*	0.94	*23.1*	0.16	*3.8*	0.3	*7.5*
Zimbabwe	6.20	5.76	*92.9*	4.07	*65.6*	4.98	*80.4*	0.16	*2.6*	0.2	*2.9*
ALL	**121.71**	**49.08**	***40.3***	**5.09**	***4.2***	**47.93**	**39.4**	**3.17**	***2.6***	**4.1**	***3.3***

Overall, these figures suggest about 40% of the regional cattle populations are at risk of ECF (*i.e*. within the vector and disease overlap zone). This broad figure conceals a great deal of variation – from zero (CAR, Cameroon, Somalia, Swaziland,) to >90% (Rwanda, Zimbabwe). Intermediate levels between 25 and 70% are estimated for many of the major cattle rearing countries: Uganda, Tanzania, Mozambique, Kenya, Burundi). These proportions equate to about 50 million cattle region wide, with the largest populations potentially at risk of ECF being in Kenya, Uganda and Tanzania.

Table 5 also provides breakdowns of the cattle numbers and proportions National totals in the overlap zones for the disease and each of the two vectors. This may prove useful if integrated with vector competence data, should it become available in future.

The "Buffalo derived risk zones" are clearly only a small part of the overall overlaps between disease, cattle and vector. Table 5 suggests that 3 to 4 million cattle may be at risk region-wide, depending on the threshold values for presence and cattle populations used to identify the cattle at risk. This represents around 3% of the regional cattle population.

A national perspective casts a slightly different light on these rather small figures: the 'worst case' assumptions using the maximum degree of overlap between buffalo risk zones and cattle result in estimates around a million (0.7 – 1.4m) cattle at risk in each of Kenya, Uganda, and Tanzania, which together account for three quarters of the region population calculated to be potentially vulnerable to Corridor Disease. Despite the numbers, however, this still equates to significantly less than 10% of the national herds.

6 Summary, Discussion and Recommendations

The risk of cattle acquiring Buffalo-derived strains of *T.parva* has been known for some considerable time. The main epidemiological and veterinary impacts are that currently available methods of treatment and prevention based on the use of the Muguga cocktail in an Infection and Treatment Method may not be effective in areas where cattle could come into contact with tick disease vectors that have fed on infected Buffalo. There is, however, some uncertainty about how widespread this risk is because the treatment has been observed to work in some areas where Buffalo and cattle coexist, but not in others.

From a veterinary perspective, there are a number of consequences. Firstly there are the direct effects on cattle mortality which are usually severe. Secondly, if the risk to cattle of Buffalo derived disease strains is extensive it may be worth investing in ways of improving the treatment delivery chains or improving the Muguga cocktail itself. Finally, if stockholders perceive the treatment to be ineffective, even if rarely, it may affect take-up rates in areas where there are no buffalo derived disease strains – which would affect ECF treatment efficacy throughout the ECF risk zone which covers a much wider area.

In this context, the objective of the current study has not been to determine the risk of treatment failure, but rather to assess the extent of the risk to cattle herds in East and Southern Africa. This involves first establishing the geographic distributions of the disease and its tick vectors, and of the Buffalo hosts, and then estimating the number of cattle that are present within the areas where the Buffalo, the vectors and the disease are all present.

This has involved a number of steps: collating the available data; filling in the gaps if needed; defining the conditions of overlap, and finally estimating the cattle numbers in the overlap zones identified.

Estimates of cattle numbers are freely available through the Gridded Livestock of the World (FAO, 2017). For the other components, extensive data searches have been carried out to acquire and compile what data are available in the public domain. For the vectors and the disease, fairly complete data are available for most of the region, though a number of gaps remain. These have been filled by well established statistical modelling techniques – specifically Random Forest and Boosted Regression Trees – supported by a very extensive set of predictor covariates including tailored environmental and climatic variables derived from long term satellite imagery time-series. to provide area-wide predictions of the probability of presence of both vectors and disease

These modelling techniques could not be used in the case of Buffalo distributions. Whilst this species is known to be restricted to areas close to open water and with at least some light forest cover they are now largely limited to protected areas and some limited and increasingly restricted areas outside them. They are not, therefore, primarily limited by environmental or habitat factors and are therefore not readily amenable to spatial modelling techniques at a subcontinental scale. Accordingly the Buffalo distributions were estimated from known densities within Protected Areas, masked by habitat suitability derived from water and forest coverages.

The spatial overlaps between disease, vectors and suitable buffalo within known distribution provided the kernel of the risk zones from which cattle could come into contact with Buffalo derived strains of the disease. Cattle are, however known to move extensively whilst grazing, and so buffer zones were defined around these kernels so that livestock up to a days' travel away (a maximum of 30km) were defined as being able to come into contact with ticks infected with the Buffalo disease strains.

The resulting estimates suggest that about 50 million cattle are within the regional ECF risk areas. Of these, and depending on the assumptions made to define the buffer zones and overlap thresholds, only 3-4 million cattle were calculated to be at risk of encountering Buffalo strains of *T.parva.* This represents around 3% of the entire cattle population, or less that 10% of those at risk of Theileriosis region-wide. Whilst this, is at first sight, a small proportion, the figure for some countries are somewhat higher – in the region of a million (0.7m – 1.4m) cattle are estimated to be at risk in each of Kenya, Uganda, and Tanzania.

These results are closely in line with numbers derived from expert opinion using entirely different sets of information (Kiara, *pers. comm*.), but the similarity is encouraging given the complete contrast in methodologies used.

Whether these numbers are large enough to justify further investment of the treatment methods depends on the cost-benefit economics of the cattle production and treatment in each area and cannot be commented upon within this study. The numbers of animals likely to be affected if the treatment fails do, however, seem to be high enough that stock holders will become aware of the problem. If this is the case then decisions as to whether to treat and warn the cattle owners of the risks, or not to treat in areas where cattle and buffalo coexist will need to me made.

To these conclusions there are, as ever, a number of caveats to be applied. In a study of this sort, which rely on literature and online data searches, and statistical rather than process based analyses to fill the gaps, it is always possible that the results are inaccurate. In this case, however, the data records used to generate the vector and disease models are sufficiently extensive to allow considerable statistical confidence in the general outcomes, with the caveat that the predictions are produced at a regional scale and so are probabilistic. Interpretation at finer scales that 30 – 50 kilometre resolutions may therefore produce misleading results.

Another significant caveat concerns seasonality and long distance movements. All the estimates produced are essentially synoptic – and so are averages. There are therefore likely to be seasonal effects that are not accounted for here. These include the fact that cattle might be present when the vectors are not active; that stock holder may avoid areas at certain parts of year, or at certain times of day in the knowledge that ticks are a particular risk during such periods. It is simply not possible for a short term study at subcontinental scales to accommodate these factors. It may, however, be more feasible to produce localised assessments with several sets of overlaps, each valid for different seasons. An example might be estimating the impact of seasonal grazing on the interactions between wild and domestic animals.

This study has led to the acquisition and collation of a large amount of geographic data – as described in Section 11. Because they are standardised these data are potentially useful for a wide range of similar analyses. The covariate data sets can be used in spatial distribution models or correlation analyses of any suitable target parameter (i.e. any disease vectors, hosts, or diseases) providing enough training data are available to calibrate the models. These data can also be used to delineate suitable habitats for any species for which the environmental or climatic limiting factors are sufficiently well understood.

The methodology used here illustrates the combination of several widely used techniques to produce distributions of components for which the overlaps can be mapped and populations within them thereby estimated. The approach can be used in many other epidemiological contexts – involving both infectious and vector-borne diseases to provide rapid assessments of the likely size of a problem caused by contact between species..

7 References

ILRI (2017) https://ilvac.net/2014/08/19/a-short-history-of-the-live-vaccine-protecting-africas-cattle-against-east-coast-fever/

Tatjana Sitt, E. Jane Poole, Gideon Ndambuki, Stephen Mwaura , Thomas Njoroge, George P. Omondi, Matthew Mutinda, Joseph Mathenge, Giles Prettejohn, W. Ivan Morrison, Philip Toye (2015). Exposure of vaccinated and naive cattle to natural challenge from buffalo-derived *Theileria parva*. International Journal for Parasitology: Parasites and Wildlife 4(2):244–251.

Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y., Hill, E.W., Rege, J.E., 2002. African pastoralism: genetic imprints of origins and migrations. Science 296 (5566), 336–339.

Josephine G. Walker, Eili Y. Klein, Simon A. Levin. Veterinary Parasitology 199 (2014) 206– 214. http://dx.doi.org/10.1016/j.vetpar.2013.11.008

Perry, B.D; (2016) . The control of East Coast fever of cattle by live parasite vaccination: A science-to-impact narrative. One Health. Jul 22;2:103-114. doi: 10.1016/j.onehlt.2016.07.002. eCollection 2016 Dec

Kiara K., Steinaa L., Svitek N., Schieck E., and Toye P. (2016). Access to livestock health interventions and products in dairy and cattle value chains in Tanzania . ILRI Research Programme of Livestock and Fish Brief 27 Dec 16

Homewood, K., Trench, P., Randall, S., Lynen, G., Bishop, B., 2006. Livestock health and socio-economic impacts of a veterinary intervention in Maasailand: infection-and-treatment vaccine against East Coast fever. Agric. Syst. 89 (2–3), 248–271

Scharlemann J., Benz D., Hay S.I., Purse B.V, Tatem A.J, Wint G.R.W, and Rogers D.J, (2008) Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data PLoS ONE 3(1): e1408. doi:10.1371/journal.pone.0001408

Kabi, F., Masembe C., Muwanika, V , Kirunda H., Negrini R. (2014). Geographic distribution of non-clinical *Theileria parva* infection among indigenous cattle populations in contrasting agro-ecological zones of Uganda: implications for control strategies. Parasites & Vectors 2014, 7:414. http://www.parasitesandvectors.com/content/7/1/414

ILRAD, (1989) Proceedings of a workshop on Theilerosis held in Nairobi, http://www.fao.org/wairdocs/ilri/x5549e/x5549e01.htm#TopOfPage

Norval, R.A., Perry, B.D., Young, A., (1992) The Epidemiology of Theileriosis in Africa. Academic Press, London

De Garcia and Serrano, F. M. H., (1971). Contribuicao para o estudo da theileriose cincerina maligna dos bovines em Angola. Acta Veterinaria, Nova Lisboa. 7, 1-8.Kalumi, M.K., Losson, B. and Saegerman, C (2011) Epidémiologie et contrôle de la theilériose bovine à Theileria parva en Afrique : une revue de la littérature Ann. Méd. Vét., 155, 88-104

Stolz, H. (2017) *Theileria parva* infections. http://www.afrivip.org/sites/default/files/Theilerioses/index.html

Lawrence, J.A., Perry, B.D. & Williamson, S.M., (2005a). East Coast fever, in: Infectious diseases of livestock, edited by Coetzer, JAW & Tustin, RC. Cape Town: Oxford University Press Southern Africa.

Lawrence, J.A., Perry, B.D. & Williamson, S.M. (2005b). Corridor disease, in: Infectious diseases of
Cumming, G. S. (1999). Host distributions do not limit the species ranges of most African ticks (Acari: Ixodida). Bulletin of Entomological Research 89, 303–327

Cumming, G. S. (1999). Host distributions do not limit the species ranges of most African ticks (Acari: Ixodida). Bulletin of Entomological Research 89, 303–327

Lynen G., Zeman P.,· Bakuname C., Di Giulio G., Mtui P., Sanka P., Jongejan F., (2007) Cattle ticks of the genera Rhipicephalus and Amblyomma of economic importance in Tanzania: distribution assessed with GIS based on an extensive Weld survey Exp Appl Acarol (2007) 43:303–319. DOI 10.1007/s10493-007-9123-9

Sungirai, M., Abatih, E. N., Moyo, D. Z., Clercq, P. D. and Madder, M. (2017), Shifts in the distribution of ixodid ticks parasitizing cattle in Zimbabwe. Med Vet Entomol, 31: 78–87. doi:10.1111/mve.12215

Spickett A. M., I. Heyne I. H., and Williams R., (2011) Survey of the livestock ticks of North West province South Africa. Onderstepoort J Vet Res, Vol 78, No 1

Gachohi J., Skilton R., Hansen R., Ngumi P., and Kitala P. (2012) Epidemiology of East Coast fever (Theileria parva infection) in Kenya: past, present and the future. Parasites & Vectors 5:194 https://doi.org/10.1186/1756-3305-5-194

Madder M., Horak I., and Stoltsz H. (2017) http://www.afrivip.org/sites/default/files/Ticks-importance/rhipicephalus.html

Bontemps S., Van Bogaert E., Defourny P., Kalogirou V. and Arino O., GlobCover 2009 – Products Description Manual", version 1.0, December 2010. Available on the ESA IONIA website (http://ionia1.esrin.esa.int/).

FAO (2017). Gridded Livestock of the World. http://www.fao.org/ag/againfo/resources/en/glw/home.html

Wint, W. and Robinson, T.P. (2007) Gridded Livestock of The World. FAO, Rome.

Cecchi, G., Wint W., Shaw A., Marletta A., Mattioli R., and Robinson, T. (2010) Geographic distribution and environmental characterization of livestock production systems in East Africa. Agriculture, Ecosystems and Environment 135 (2010) 98–110

Sere C. and Steinfeld H. (1996) World Livestock Production Systems. Current status, issues and trends. FAO Animal Production and Health Paper 127. FAO Rome

Robinson T.P., Thornton P.K., Franceschini, G., Kruska, R.L., Chiozza, F., Notenbaert, A., Cecchi, G., Herrero, M., Epprecht, M., Fritz, S., You, L., Conchedda, G. & See, L. (2011) Global livestock production systems . Rome, Food and Agriculture Organization of the United Nations (FAO) and International Livestock Research Institute (ILRI), 152 pp

East, R. (1999) African Antelope Database 1998. IUCN/SSC Antelope Specialist Group. IUCN Species Survival Commission Occasional Paper No 21. IUCN, Gland, Switzerland and Cambridge, UK.

Cornélis D., Melletti D., Korte D., Ryan S.J., Mirabile M., Prin T., and Prins H.H.T. (2014), African Buffalo, *Syncerus caffer* (Sparrman, 1779). Chapter 20 in: Ecology, Evoloution and Behaviour of Wild Cattle: Implications for Conservation. Ed. M. Melletti and J. Burton. Cambridge University Press

Robinson T.P. and Seimbieda, J. (2011) Mapping African buffalo distributions, in relation to livestock disease risk. Unpublished Study for the Pro Poor Livestock Policy Initaiative, FAO Rome

IUCN (2017) Mammal distribution database http://www.iucnredlist.org/download_spatial_data

Furstenberg D., (2009) Buffalo distribution map. Published Online Only. http://www.wildliferanching.com/content/cape-buffalo-syncerus-caffer

IEA (1998), AMD African Mammals Databank - A Databank for the Conservation and Management of the African Mammals Vol 1 and 2. Report to the Directorate-General for Development (DGVIII/A/1) of the European Commission. Project No. B7-6200/94-15/VIII/ENV. Bruxelles 1998. pp 1174.

Prins H.H.T., (1996)","Ecology and behavior of the African buffalo. Social inequality and decision making.","Chapman & Hall, London. Wildlife Ecology and Behaviour Series: n° 1, 293 pp

Kingdon J., (1997) The Kingdon field guide to African Mammals.","Academic Press, London and New York: Natural World, 464 pp

ESA (2017) Land Cover CCI 2015 100m resolution .
http://maps.elie.ucl.ac.be/CCI/viewer/download.php

ESA (2016). - S2 prototype Land Cover 20m map of Africa 2016
http://2016africalandcover20m.esrin.esa.int/

VITO (2017) Small Water bodies. http://land.copernicus.eu/global/products/wb

USGS (2006) Hydrosheds database (Lehner B., Verdin K., and Jarvis A.)
https://hydrosheds.cr.usgs.gov/index.php

Mbizeni S., Potgieter F. T., Troskie C., Mans B. J., Penzhorn B. L, Latif A. A. Ticks and Tick-borne Diseases 4 (2013) 227– 234 http://dx.doi.org/10.1016/j.ttbdis.2012.11.005

8 Additional Figures

Figure 11: Parasite BRT Model Distributions: a) Extended Distributions (left), b) Consensus Distributions (right)

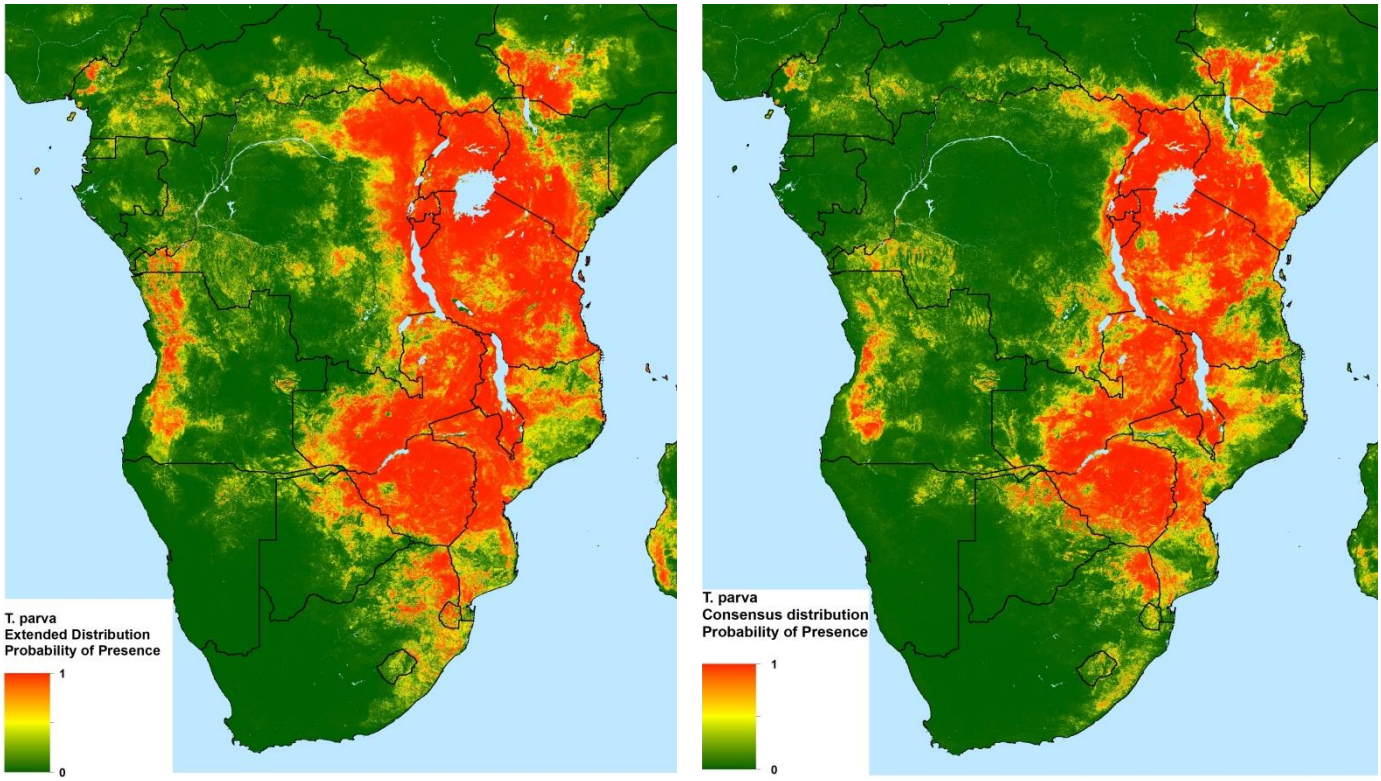


Figure 12: Predicted Distribution, either Vector

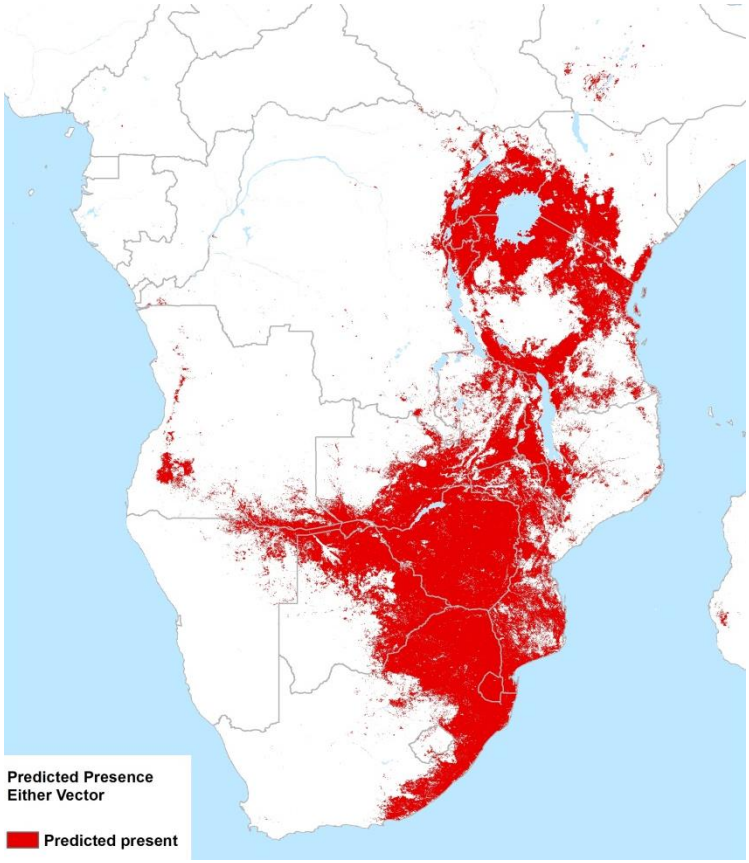


Figure 13: a) Overlap between Disease and either Vector; b) Cattle in ECF Overlap

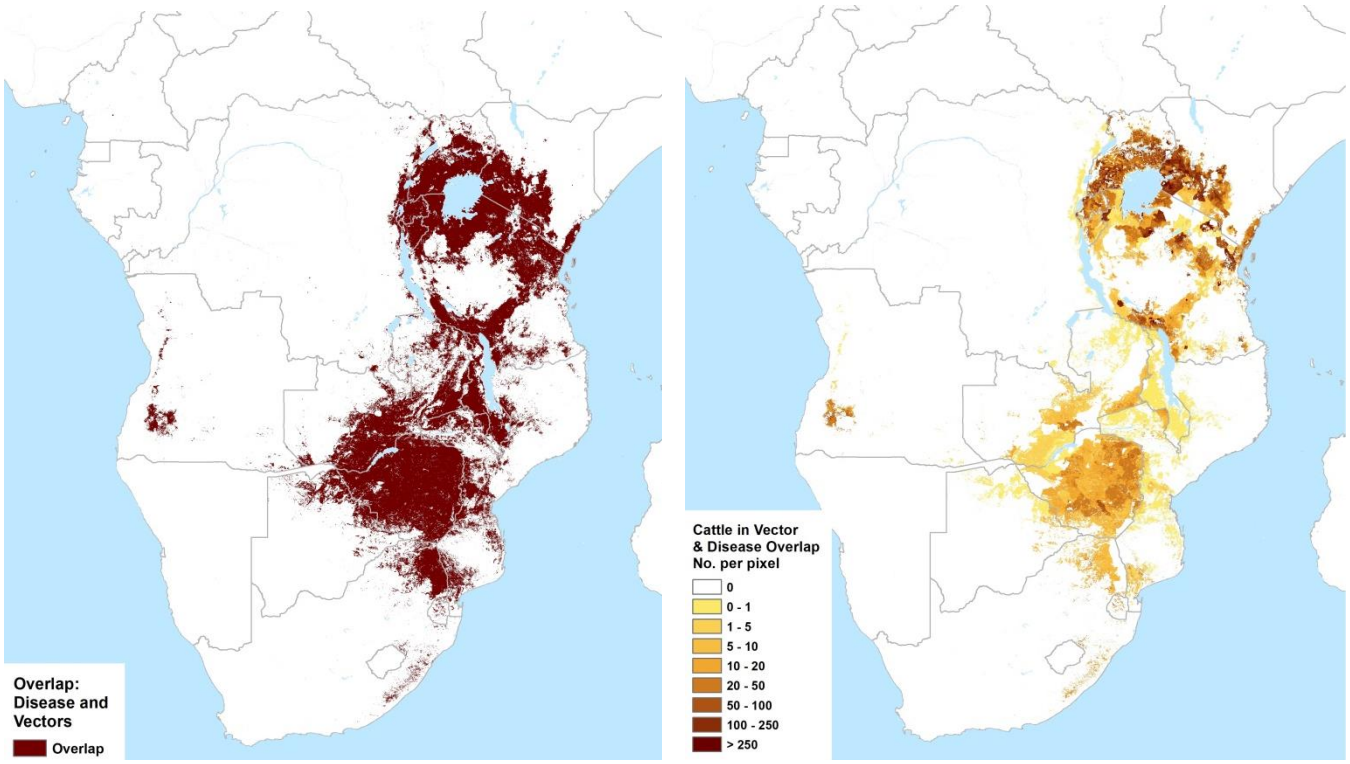
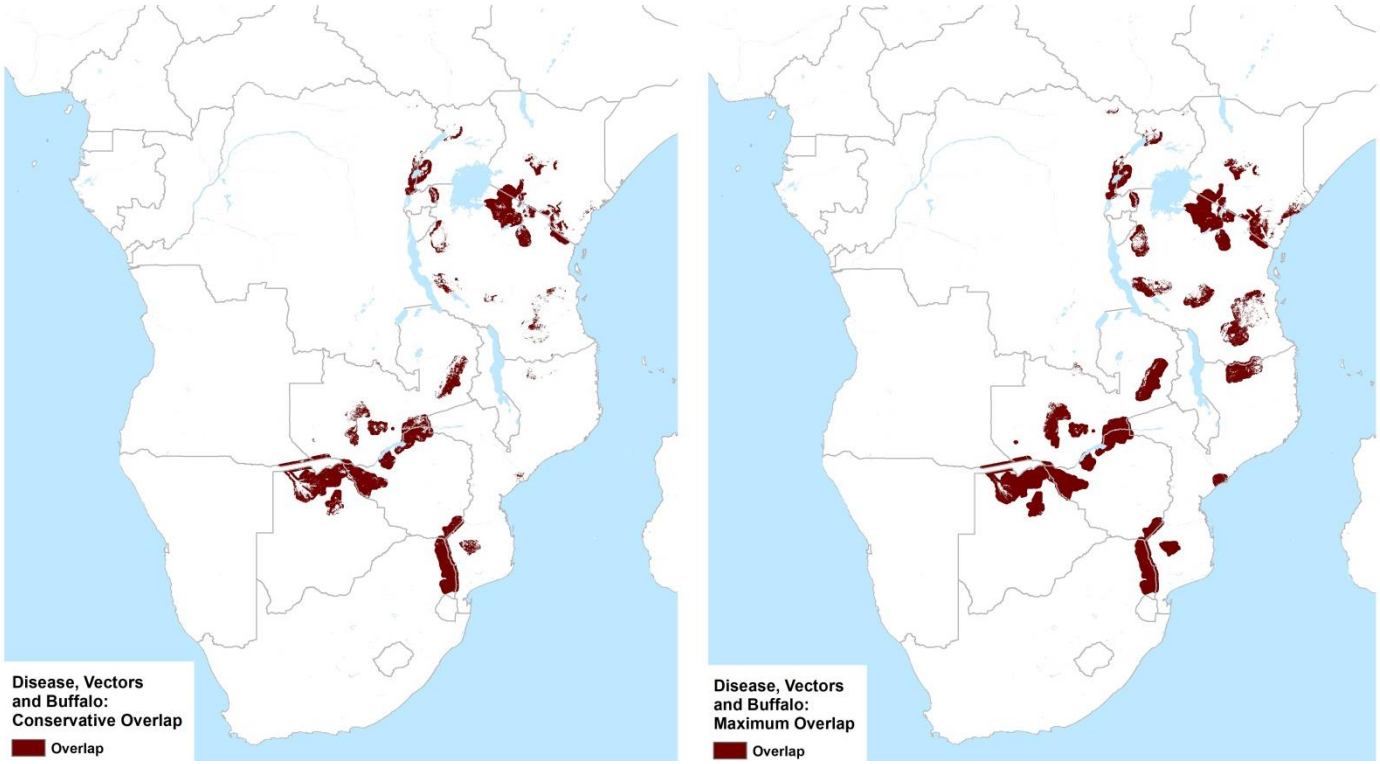


Figure 14: Buffalo derived risk zones a) Conservative; b) Maximum



9 Terms of Reference

The scanned Terms of Reference are shown below.

TERMS OF REFERENCE

CONSULTANT NAME Geoffrey Roger William Wint
ASSIGNMENT/PROJECT NAME: Cluster 2 and Cluster4-Risk assessment and Animal Health Delivery
COUNTRY OF ASSIGNMENT: UK

1) GENERAL BACKGROUND

The infection and treatment method (ITM) of immunisation against East Coast fever (ECF) has more than 95% efficacy against cattle derived T.parva infection. The evidence of the performance of the vaccine against Buffalo-derived T.parva is mixed. In some situations vaccination failed to protect cattle against buffalo-derived T.parva infection. In other areas, the vaccine appears to work perfectly even when cattle and buffalo graze together.

It is not clear if the reason for this difference is due to buffalo-derived parasites breaking through the immunity engendered by the Muguga cocktail strains and where the vaccine works is because somehow the buffalo derive T.parva stocks are protected by the vaccines stocks. It is now known that buffalo derived parasites are more polymorphic than cattle derived parasites. It might also be that in the areas where the vaccine works, the proportion of buffalo derived parasites compared to cattle derived parasites is so low that the vaccine is able to protect against the cattle parasites which happen to be in the majority.

These issues require investigation to get a better understanding so that the vaccine could be improved. In the meantime, in order to protect the reputation of the vaccine, it is safer not to vaccinate in areas where cattle are known to interact with buffalo. Unfortunately these areas are not clearly known. Besides the number cattle at risk from the buffalo-derived parasites is not known to decide if it is worthwhile to invest in protecting those cattle.

2) OBJECTIVE OF THE ASSIGNMENT

This study hopes to map the areas considered to be at risk and therefore to be avoided for ITM with the Muguga cocktail.

3) SCOPE OF WORK

- Assemble all available data layers in a common format for Africa (cattle, cattle production systems, buffalo, T. parva and relevant tick species (R.appendiculatus and R. zambeziensis) - 15 days
- Model the spatial distribution of any of these factors for which continuous, geographic estimates are not available or need updating. - 10 days
- Review the literature in relation to risk of cattle infection from buffalo derived T. parva and the current position with regard to protection buffalo derived parasites by the Muguga cocktail ITM vaccine. - 5 days
- Combine these maps in such a way as to estimate risk of from buffalo derived T. parva. - 5 days
- Assemble a GIS database with all input data and results. - 5 days
- Write a report and draft a manuscript for submission to a peer-reviewed journal presenting the study. - 10 days

4) DURATION OF ASSIGNMENT, DUTY STATION AND EXPECTED PLACES OF TRAVEL

Duration: 2nd August - 30 November 2017
Maximum working days: 50 days
Duty Station: Oxford - UK
Expected places of travel: Nairobi

10 Mission to ILRI Activities

Date/Period	Meeting topic/activity	Personnel
3rd December: Evening	Arrival Nairobi	
4th December: Morning Afternoon:	Familiarisation Presentation of Initial results and potential issues	Dr Henry Kiara
5th December: Morning: Morning: Afternoon:	Data delivery and required data formats Discuss technical approach, condition of spread from buffalo to cattle Risk to cattle of Buffalo derived disease in Kenya	Dr Catherine Pfeifer (ILRI) Dr Phil Toye (ILRI) Dr S. Ndungu and Colleagues (Muguga Veterinary Research Institute)
6th December Morning: Morning: Afternoon:	Availability of Buffalo and Livestock data for Kenya Context of using study results, allied studies Presentation of Study and draft results to ILRI staff	Dr Patrick W. Wagute, Director, Directorate of Resource Surveys and Remote Sensing Dr Vish Nene (ILRI)
7th December Morning: Afternoon	Presentation of Study to Directorate of Veterinary Services, Kabete Revision of analyses to accommodate visit findings	Dr Harry Oyas (Deputy Director Veterinary Services) & Colleagues
8th December Morning: Afternoon: Evening:	Briefing Senior ILRI Staff Debriefing and wrap-up Depart Nairobi	Dr Dieter Schillinger Dr Henry Kiara

11 Data Catalogue

All data have been made available to ILRI via posts on Google drive. File format details, Input and output files and covariate files are listed in the following sections and in an Excel spreadsheet ILRECFDATALIST.XLS supplied with this report.

11.1 README notes

Details of the file formats are as follows:

Unless otherwise stated the properties of all Raster images are as follows:

Pixel size: 0.00833333 x 0.00833333 degrees

Rows: 6000

Columns 8520

Projection: GCS_WGS_1984

Datum D_WGS_1984

Image Format: geotiff

Image Component files tif, tfw

Top Coordinate (degrees) 14

Left Coordinate (degrees) -18

Right Coordinate (degrees) 53

Bottom Coordinate (degrees) -36

Unless otherwise stated all vector files are ESRI shapes

Projection: GCS_WGS_1984

Datum D_WGS_1984

Vector Component files shp, shx, sbn (prj)

The maps used in the report Figures can be displayed in ARcMAP using the following document files:

ECFDELIVERY93.MXD ArcMAP V9x document file

ECFDELIVERY10.MXD ArcMAP V10x document file

11.2 Input and Output files

Map Description	Report Figure	Delivery Filenames
Bounding coordinates for this ECF study	Figure 1	\ILRIECF\Delivery\ecfextent.tif & .jpg
Land Water Mask for ECF Extent	All figures	\ILRIECF\Delivery\eclandmask.tif & .jpg
Country Boundaries	All figures	\ILRIECF\Delivery\af_country.tif & .jpg
ECF Countries	Figure 2a	\ILRIECF\Delivery\ECFCOUNTRIES.shp & .jpg
Distribution of Theileriosis		
Published maps	Figure 2a	\ILRIECF\Delivery\tparavadatasets.shp & .jpg
Presence and absence points assigned	Figure 2b	\ILRIECF\Delivery\tparvaPApoints.shp & .jpg
Predicted Parasite Distribution, Consensus Distribution		
Extended BRT Model	Figure 11	\ILRIECF\Delivery\TPExtendedBRTModel.tif & .jpg
Consensus BRT model	Figure 11	\ILRIECF\Delivery\TPConsensusBRTModel.tif & .jpg
Consensus RF Zoned model	Not Illustrated	\ILRIECF\Delivery\TPConsensusRFZModel.tif & .jpg
Consensus Ensemble Model	Figure 3	\ILRIECF\Delivery\TPConsensusEnsembleModel.tif & .jpg
Recorded Vector Distributions		
R.appendiculatus	Figure 4a	\ILRIECF\Delivery\rappendprespts.shp & .jpg, rappendabspts.shp
R. zambeziensis	Figure 4b	\ILRIECF\Delivery\rzambprespts.shp & .jpg
Masked Suitable Habitats		
R.appendiculatus	Not Illustrated	\ILRIECF\Delivery\RAppMaskedUnsuitability.tif & .jpg
R. zambeziensis	Not Illustrated	\ILRIECF\Delivery\RZamMaskedUnsuitability.tif & .jpg
Presence and Absence locations		
R.appendiculatus	Figure 5a	\ILRIECF\Delivery\rappendpresabspts.shp & .jpg
R. zambeziensis	Figure 5b	\ILRIECF\Delivery\rzambpresabspts.shp & .jpg
Modelled Probability of Individual Vector Presence		
R.appendiculatus BRT Model	Not Illustrated	\ILRIECF\Delivery\RAppBRTModel.tif & .jpg
R.appendiculatus Zoned RF Model	Not Illustrated	\ILRIECF\Delivery\RAppRFZonedModel.tif & .jpg
R.zambeziensis BRT Model	Not Illustrated	\ILRIECF\Delivery\RZambBRTModel.tif & .jpg
R.zambeziensis Zoned RF Model	Not Illustrated	\ILRIECF\Delivery\RZambRFZonedModel.tif & .jpg
Ensemble Vector Models		
R. appendiculatus	Figure 6a	\ILRIECF\Delivery\RAppEnsembleModel.tif & .jpg
R. zambeziensis	Figure 6b	\ILRIECF\Delivery\RZambEnsembleModel.tif & .jpg
Either Vector > 50% probability	Figure 12	\ILRIECF\Delivery\EitherVectorGT50PC.tif & .jpg
Host Distributions		
Cattle numbers per pixel corrected to FAO 2014 Totals	Figure 7a	\ILRIECF\Delivery\FAOGLWCattleperPixel.tif & .jpg
Livestock Production Zones	Figure 7b	\ILRIECF\Delivery\LivestockZonesS&S2011.tif & .jpg
Available Buffalo Distribution datasets	Figure 8a	\ILRIECF\Delivery\AvailableBuffaloDatasets.shp & .jpg
Selected Buffalo Densities Used	Figure 8b	\ILRIECF\Delivery\SelectedBuffaloDensity.shp & .jpg
Buffalo Suitable Land Cover/Land Use	Not Illustrated	\ILRIECF\Delivery\BuffaloSuitabilityLandUse.tif & .jpg
Buffalo Suitability within Range of Water	Not Illustrated	\ILRIECF\Delivery\BuffaloSuitabilityWater.tif & .jpg
Suitable Buffalo Habitat	Figure 9a	\ILRIECF\Delivery\BuffaloSuitability.tif & .jpg
Estimated range and distribution of Buffalo within suitable habitats	Figure 9b	\ILRIECF\Delivery\BuffaloDensitySuitable.tif & .jpg
Buffalo, Buffered Suitable Areas		\ILRIECF\Delivery\BuffaloBufferedSuitableAreas.tif & .jpg
Overlap ECF and Vectors		

Map Description	Report Figure	Delivery Filenames
R.appendiculatus	Not Illustrated	\ILRIECF\Delivery\OverlapDisease&RApp.tif & .jpg
R. zambeziensis	Not Illustrated	\ILRIECF\Delivery\OverlapDisease&RZam.tif & .jpg
Either Vector	Figure 13a	\ILRIECF\Delivery\OverlapDisease&Vectors.tif & .jpg
Buffalo derived risk zones		
Conservative Overlap	Figure 14a	\ILRIECF\Delivery\DisVecBufConsOverlap.tif & .jpg
Maximum Overlap	Figure 14b	\ILRIECF\Delivery\DisVecBufMaxOverlap.tif & .jpg
Cattle in Overlap Zones		
Conservative Overlap	Figure 10a	\ILRIECF\Delivery\CattleConsBufOverlap.tif & .jpg
Maximum Overlap	Figure 10b	\ILRIECF\Delivery\CattleMaxBufOverlap.tif & .jpg
Disease and either Vector	Figure 13b	\ILRIECF\Delivery\CattleDisVecOverlap.tif & .jpg

11.3 Covariate files

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\basemaps	AF_COUNTRY	Country Boundaries			
\ILRIECF\basemaps	ECFCOUNTRIES.shp	Countries within ECF zone	derived from AF_COUNTRY above		
\ILRIECF\basemaps	ECDISTWAT.tif	Distance (degrees) to Remotely Sensed water			Surface water from ESA CCI 20m and 100m Land Cover maps
\ILRIECF\basemaps	ecgmtedp5c.tif	Minimum Elevation (m) PLUS 500	Global Multi-resolution Terrain Elevation Data 2010	https://topotools.cr.usgs.gov/gmted_viewer/	500 added to global coverage to remove negative values below sea level
\ILRIECF\basemaps	ECNOTPARKS.tif	Un protected areas			Inverse of ECPARKSWW.tif
\ILRIECF\basemaps	ECPARKSWW.tif	National Parks and Game Reserves 2016	IUCN World Database on Protected Areas	https://www.iucn.org/theme/protected-areas/our-work/quality-and-effectiveness/world-database-protected-areas-wdpa	All areas with following criteria extracted from regional shape file and converted to geotiff: DESIG_ENG" LIKE '%Game%' OR "DESIG_ENG" LIKE '%Park%' OR "NAME" LIKE '%Game%' OR "NAME" LIKE 'Park' OR "IUCN_CAT" LIKE '%I%' OR "IUCN_CAT" = 'V'
\ILRIECF\basemaps	ecserestei20i11zones.tif	Livestock Zones	FAO,	https://cgspace.cgiar.org/bitstream/handle/10568/10537/faoglobalLivestock.pdf	
\ILRIECF\basemaps	ecmask00833zero	Binary Land water Mask template for this ECF study area			
\ILRIECF\basemaps	ECFEXTENT.tif	Bounding coordinates for this ECF study			
\ILRIECF\covariates\lc	ECEEBARE.tif	Percentage bare ground	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEDCBD.tif	Percentage decidupus broadleaved trees	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEEVGBD.tif	Percentage evergereen broadleaved trees	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEEVGND.tif	Percentage evergreen needleleaved gtrees	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEFLOOD.tif	Percentage flooded or irrigated vegetation	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEHERB.tif	Percentage herbaceous covers	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEMANAG.tif	Percentage managed land (e.g. cropping)	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEOTHTR.tif	Percentage other	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEESHRUB.tif	Percentage shrub cover	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\covariates\lc	ECEESNOW.tif	Percentage snow	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEURB.tif	Percentage urban	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECEEWATER.tif	Percentage open water	Tuanmu, M.-N. and W. Jetz. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modeling. Global Ecology and Biogeography 23(9): 1031-1045.	http://www.earthenv.org/landcover.html	Part of the 1km resolution consensus Land Cover datasets
\ILRIECF\covariates\lc	ECGBC100PR.tif	Proportion Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	0.008333 deg resolution calculation Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC110PR.tif	Proportion Mosaic forest or shrubland (50-70%) / grassland (20-50%)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC11PR.tif	Proportion Post-flooding or irrigated croplands (or aquatic)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC120PR.tif	Proportion Mosaic grassland (50-70%) / forest or shrubland (20-50%)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC130PR.tif	Proportion Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC140PR.tif	Proportion Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC14PR.tif	Proportion Rainfed croplands	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC150PR.tif	Proportion Sparse (<15%) vegetation	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC160PR.tif	Proportion Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily) - Fresh or brackish water	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC170PR.tif	Proportion Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\covariates\lc	ECGBC180PR.tif	Proportion Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC190PR.tif	Proportion Artificial surfaces and associated areas (Urban areas >50%)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC200PR.tif	Proportion Bare areas	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC20PR.tif	Proportion Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC210PR.tif	Proportion Water bodies	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC220PR.tif	Proportion Permanent snow and ice	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC230PR.tif	Proportion No data (burnt areas, clouds,…)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC30PR.tif	Proportion Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC40PR.tif	Proportion Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC50PR.tif	Proportion Closed (>40%) broadleaved deciduous forest (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC60PR.tif	Proportion Open (15-40%) broadleaved deciduous forest/woodland (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC70PR.tif	Proportion Closed (>40%) needleleaved evergreen forest (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\lc	ECGBC90PR.tif	Proportion Open (15-40%) needleleaved deciduous or evergreen forest (>5m)	Sophie Bontemps, Pierre Defourny, Eric V. Bogaert, et al. GLOBCOVER 2009 - Products description and validation report (February 2011)	http://due.esrin.esa.int/page_globcover.php	Derived from 300m resolution Land Cover ESA GLOBCOVER 2009 dataset
\ILRIECF\covariates\modis	scharlemannetal2008MODIS.pdf	Reference describing MODIS Fouurier Transformed Covariates			

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\covariates\modis	ModisScaling.doc	Document setting out scaleing of MODS covariate Values			
\ILRIECF\covariates\modis	EC011503A0.tif	Channel Three Temperature, Fourier Mean	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503A1.tif	Channel Three Temperature, Fourier Component 1 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503A2.tif	Channel Three Temperature, Fourier Component 2 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503A3.tif	Channel Three Temperature, Fourier Component 3 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503D1.tif	Channel Three Temperature, Fourier Component 1 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503D2.tif	Channel Three Temperature, Fourier Component 2 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503D3.tif	Channel Three Temperature, Fourier Component 3 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503MN.tif	Channel Three Temperature, Fourier Minimum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503MX.tif	Channel Three Temperature, Maximum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503P1.tif	Channel Three Temperature, Fourier Component 1 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503P2.tif	Channel Three Temperature, Fourier Component 2 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011503P3.tif	Channel Three Temperature, Fourier Component 3 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507A0.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Mean	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507A1.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 1 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507A2.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 2 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507A3.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 3 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507D1.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 1 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507D2.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 2 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507D3.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 3 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507MN.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Minimum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507MX.tif	Night-time Land Surface Enhanced Vegetation Index, Maximum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507P1.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 1 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507P2.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 2 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011507P3.tif	Night-time Land Surface Enhanced Vegetation Index, Fourier Component 3 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508A0.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Mean	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\covariates\modis	EC011508A1.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 1 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508A2.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 2 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508A3.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 3 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508D1.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 1 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508D2.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 2 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508D3.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 3 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508MN.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Minimum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508MX.tif	Day-time Land Surface Enhanced Vegetation Index, Maximum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508P1.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 1 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508P2.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 2 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011508P3.tif	Day-time Land Surface Enhanced Vegetation Index, Fourier Component 3 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514A0.tif	Normalised Difference Vegetation Index, Fourier Mean	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514A1.tif	Normalised Difference Vegetation Index, Fourier Component 1 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514A2.tif	Normalised Difference Vegetation Index, Fourier Component 2 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514A3.tif	Normalised Difference Vegetation Index, Fourier Component 3 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514D1.tif	Normalised Difference Vegetation Index, Fourier Component 1 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514D2.tif	Normalised Difference Vegetation Index, Fourier Component 2 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514D3.tif	Normalised Difference Vegetation Index, Fourier Component 3 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514MN.tif	Normalised Difference Vegetation Index, Fourier Minimum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514MX.tif	Normalised Difference Vegetation Index, Maximum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514P1.tif	Normalised Difference Vegetation Index, Fourier Component 1 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514P2.tif	Normalised Difference Vegetation Index, Fourier Component 2 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011514P3.tif	Normalised Difference Vegetation Index, Fourier Component 3 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515A0.tif	Enhanced Vegetation Index, Fourier Mean	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515A1.tif	Enhanced Vegetation Index, Fourier Component 1 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515A2.tif	Enhanced Vegetation Index, Fourier Component 2 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery

Folder	Filename	Description	Source/Reference/Documentation supplied	URL	Comments
\ILRIECF\covariates\modis	EC011515A3.tif	Enhanced Vegetation Index, Fourier Component 3 Amplitude	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515D1.tif	Enhanced Vegetation Index, Fourier Component 1 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515D2.tif	Enhanced Vegetation Index, Fourier Component 2 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515D3.tif	Enhanced Vegetation Index, Fourier Component 3 % Total Variation	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515MN.tif	Enhanced Vegetation Index, Fourier Minimum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515MX.tif	Enhanced Vegetation Index, Maximum	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515P1.tif	Enhanced Vegetation Index, Fourier Component 1 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515P2.tif	Enhanced Vegetation Index, Fourier Component 2 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	EC011515P3.tif	Enhanced Vegetation Index, Fourier Component 3 Phase	scharlemannetal2008MODIS.pdf		Fourier Process datasets derived from 2000-2015 timeseries of V5 imagery
\ILRIECF\covariates\modis	ECRHJSMI15.tif	Rdelative Humidity, Jun to September 2015	Report		Extracted from dataset produced by P. Jones, for IDAMS FP7 Project
\ILRIECF\covariates\modis	ECRHMIN15.tif	Minimum Relative Humidity, 2015	Report		Extracted from dataset produced by P. Jones, for IDAMS FP7 Project
\ILRIECF\covariates\modis	ECHFOTTV2.tif	Huam Footprint, Version 2	http://sedac.ciesin.columbia.edu/downloads/maps/wildareas-v2/wildareas-v2-human-footprint-geographic/hfp-world.pdf	sedac.ciesin.columbia.edu/data/set/wildareas-v2-human-footprint-geographic/maps	Index of Human footprint on Environment
\ILRIECF\covariates\modis	ECWPPDN15A.tif	WORLDPOP Human Population Density, 2015		www.worldpop.org.uk/	
\ILRIECF\covariates\modis	ECTSAT0116.tif	Tamsat monthly precipitation (mm), 2016, January		https://www.tamsat.org.uk	
\ILRIECF\covariates\modis	ECTSAT0216.tif	Tamsat monthly precipitation (mm), 2016 February		https://www.tamsat.org.uk	
\ILRIECF\covariates\pop	ECTSAT0316.tif	Tamsat monthly precipitation (mm), 2017 March		https://www.tamsat.org.uk	
\ILRIECF\covariates\pop	ECTSAT0416.tif	Tamsat monthly precipitation (mm), 2016, April		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT0516.tif	Tamsat monthly precipitation (mm), 2016 may		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT0616.tif	Tamsat monthly precipitation (mm), 2017 June		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT0716.tif	Tamsat monthly precipitation (mm), 2016, July		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT0816.tif	Tamsat monthly precipitation (mm), 2016 August		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT0916.tif	Tamsat monthly precipitation (mm), 2017 September		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT1016.tif	Tamsat monthly precipitation (mm), 2016, October		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT1116.tif	Tamsat monthly precipitation (mm), 2016 November		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT1216.tif	Tamsat monthly precipitation (mm), 2016 Total		https://www.tamsat.org.uk	
\ILRIECF\covariates\prec	ECTSAT2016.tif	Tamsat monthly precipitation (mm), 2016 December		https://www.tamsat.org.uk	Derived from monthly totals
\ILRIECF\covariates\prec	ECSWB16MN.tif	bodies, Proportion Year present		http://land.copernicus.eu/global/products/wb	Derived from dekadal presence of small water bodies -